

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

4 protein - protein search, using sw model

in on: September 29, 2003, 08:01:08 ; Search time 44 seconds
(without alignments)

1111.086 Million cell updates/sec

title: US-10-088-548-2

effect score: 1786

sequence: 1 KSPVVDYCHGDSYRGIS.....YTMNPKLFYCDIPLCASS 308

scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 1107863 seqs, 158726573 residues

total number of hits satisfying chosen parameters: 944443

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A: Geneseq.19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Query	Score	Match	Length	ID	Description
1	1786	100.0	308	22	AA72944	Human angiogenesis
2	913	51.1	189	21	AA01918	Human plasminogen
3	913	51.1	192	21	AA01919	Human plasminogen
4	744	41.7	280	20	AA702108	A multifunctional
5	744	41.7	297	20	AA702103	A multifunctional
6	716	40.1	250	17	AAW07570	Bovine kringle 1-3
7	715	40.0	250	17	AAW07569	Porcine kringle 1-
8	713	39.9	268	20	AA702109	A multifunctional
9	713	39.9	285	20	AA702102	A multifunctional

10	712	39.9	260	24	AA79748	Human plasminogen
11	712	39.9	274	21	AA01910	Human plasminogen
12	711	39.8	260	21	AA26486	Deglycosylated kri
13	711	39.8	260	22	AA01209	Human Angiostatin
14	707	39.6	254	21	AA53869	Amino acid sequenc
15	704	39.4	250	17	AAW07568	Rhesus kringle 1-3
16	703	39.4	250	17	AAW07567	Human kringle 1-3.
17	700	39.2	250	17	AAW07566	Murine kringle 1-3
18	669	37.5	269	22	AA84518	Amino acid sequenc
19	585	32.8	210	20	AAV06192	Anti-angiogenic an
20	539	30.2	91	22	AA72945	Human apolipoprote
21	538	30.1	168	17	AAW07561	Murine kringle 2-3
22	531	29.7	168	17	AAW07564	Porcine kringle 2-
23	524	29.3	130	23	AB77206	Fusion protein Kri
24	523	29.3	89	22	AA72946	Human apolipoprote
25	519	29.1	91	20	AA42303	Human apolipoprote
26	516	28.9	120	15	AA847044	Lipoprotein. Synt
27	516	28.9	184	20	AAV02107	A multifunctional
28	516	28.9	201	20	AAV02104	A multifunctional
29	515	28.8	168	17	AAW07563	Rhesus kringle 2-3
30	515	28.8	172	22	AAW05022	Human recombinant
31	515	28.8	175	21	AA01911	Human plasminogen
32	514	28.8	168	17	AAW07565	Bovine kringle 2-3
33	511	28.6	168	17	AAW07562	Human kringle 2-3.
34	508	28.4	86	22	AA72947	Human apolipoprote
35	503	28.2	160	17	AAW07575	Bovine kringle 1-2
36	499	27.9	160	17	AAW07574	Porcine kringle 1-
37	490	27.4	160	17	AAW07573	Rhesus kringle 1-2
38	489	27.4	160	17	AAW07572	Human kringle 1-2
39	486	27.2	160	17	AAW07571	Murine kringle 1-2
40	481.5	27.0	290	13	AA22464	Truncated hepatocy
41	481.5	27.0	290	14	AA30997	34k HGF variant H
42	481.5	27.0	290	18	AAW23712	34 Kilodalton hepa
43	473	26.5	101	18	AAW34286	Human kringle 5 pe
44	473	26.5	101	21	AA01890	Human plasminogen
45	473	26.5	101	21	AA01912	Human plasminogen

ALIGNMENTS

RESULT 1
AA72944
ID: AA72944 standard; Protein; 308 AA.
XX: AA72944;
AC: AA72944;
XX: AA72944;
DT: 13-JUN-2001 (first entry)
DE: Human angiogenesis inhibitor, LK68 protein.
KW: Human; angiogenesis inhibitor; LK68; apolipoprotein(a) kringle domain;
KW: angiogenesis-mediated disease; cancer; rheumatoid arthritis; therapy;
KW: cytostatic; antirheumatic; antiarthritic; antipsoriatic; psoriasis;
KW: ocular angiogenic disease; endothelial cell proliferation; tumour;
KW: cell migration.
OS: Homo sapiens.
XX: Homo sapiens.
FH: Key
FT: Domain
FT: Location/Qualifiers
FT: 1..91
FT: /label= KIV36
FT: /note= "Apolipoprotein(a) kringle domain IV36, LK6
FT: protein"
FT: 119..207
FT: /label= KIV37
FT: /note= "Apolipoprotein(a) kringle domain IV37, LK7
FT: protein"
FT: 223..308
FT: /label= KV38
FT: /note= "Apolipoprotein(a) kringle domain V38, LK8
FT: protein"
XX: XX

W02000119868-A1.
22-MAR-2001.
15-SEP-1999; 99WO-KR00554.
15-SEP-1999; 99WO-KR00554.
(MOGA-) MOGAM BIOTECHNOLOGY RES INST.
Chang J, Kim JS, Park EJ, Yum J, Chung S;
WPI: 2001-244787/25.
N-PSDB; AAD03255.
Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated diseases, e.g. cancer and rheumatoid arthritis, has human apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid sequence -
Claim 4; Page 42-44; 50pp; English.
The present sequence is human angiogenesis inhibitor, LK68 protein.
LK68 protein contains the amino acid sequences of human apolipoprotein(a) kringle domains IV36 (LK6 protein), IV37 (LK7 protein) and V38 (LK8 protein). LK68, LK6, LK7 and LK8 are inhibitors of endothelial cell proliferation, cell migration and normal development of capillaries in the chick embryo chorioallantoic membrane (CAM). LK68 protein, its single kringles or their functional equivalents, are useful for treating angiogenesis-mediated diseases, such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic disease in animals or humans. LK68 is useful as an anticancer agent and also for inhibiting primary tumour growth.
Sequence 308 AA;
Query Match 100.0%; Score 1786; DB 22; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KSPVODCYHGDGRSYRIGISSTVTGRTQCSWSSMPHWHQRTPEPNYAGLTENYCRNP 60
1 KSPVODCYHGDGRSYRIGISSTVTGRTQCSWSSMPHWHQRTPEPNYAGLTENYCRNP 60
61 DSGKQPCWYTDPCVRWEYCNLFQCSSETSGVLEPTVPVPSMEARSAAPTETPVVR 120
61 DSGKQPCWYTDPCVRWEYCNLFQCSSETSGVLEPTVPVPSMEARSAAPTETPVVR 120
121 QCYHNGQSYRGTSTVTGRTQCSWSSMPHWHQRTPEPNYAGLTENYCRNPADTGP 180
121 QCYHNGQSYRGTSTVTGRTQCSWSSMPHWHQRTPEPNYAGLTENYCRNPADTGP 180
181 WCFTTDPSPRWEYCNLFQCSSETSGVLEPTVPVPSMEARSAAPTETPVVR 240
181 WCFTTDPSPRWEYCNLFQCSSETSGVLEPTVPVPSMEARSAAPTETPVVR 240
241 TVTGTQCEWAAQEPHRSHTFIPGTNKGAGLEKNCYCRNPDDINGPWCYTMNPKLFYD 300
241 TVTGTQCEWAAQEPHRSHTFIPGTNKGAGLEKNCYCRNPDDINGPWCYTMNPKLFYD 300
301 DIPLCASS 308
301 DIPLCASS 308
AAB01918 standard; Protein; 189 AA.
AAB01918;
18-SEP-2000 (first entry)

DE XX Human plasminogen kringle 4-5 (Val355-Ala543).
XX XX Plasminogen; human; kringle domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antiproliferative; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX Homo sapiens.
OS US6057122-A.
PN 02-MAY-2000.
XX 05-MAY-1997; 97US-0851350.
XX 03-MAY-1996; 96US-0643219.
PR 03-APR-1997; 97US-0832087.
XX (ABBO) ABBOTT LAB.
XX Davidson DJ;
XX WPI: 2000-349573/30.
XX Preparation of Kringle five peptide fragment for treating various disorders such as angiogenic, ocular, skin diseases and cancer, involves mixing mammalian plasminogen and elastase followed by incubation and isolation -
XX Example 17; Page -; 48pp; English.
XX The invention relates to a method of preparing plasminogen kringle 5 peptide fragments. The method comprises mixing mammalian plasminogen and elastase in the ratio 1:100-1:300, followed by incubating and isolating the fragment. The kringle 5 peptides are inhibitors of angiogenesis and endothelial cell proliferation and migration. The peptides are useful for treating angiogenic diseases, primary and metastatic solid tumours and carcinomas of various organs such as breast, genital tract, endocrine glands, skin, tumours of the brain and eyes and solid tumours arising from haematopoietic malignancies such as leukaemias and lymphomas. They are also used for the prophylaxis of various autoimmune diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber syndrome), diseases caused by excessive or abnormal stimulation of endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases which have angiogenesis as a pathologic consequence (e.g., cat scratch disease and ulcers). The peptides are also useful as a birth control agent which inhibits ovulation and establishment of the placenta.
XX Sequences AAB01906-B01919 represent fragments of human plasminogen used in an exemplification of the invention.
XX Note: This sequence is not shown in the specification, but is derived from the full length human plasminogen sequence (AAB01887) shown in figure 1.
XX Sequence 189 AA;
QY Query Match 51.1%; Score 913; DB 21; Length 189;
Best Local Similarity 79.4%; Pred. No. 1.3e-51;
Matches 150; Conservative 17; Mismatches 22; Indels 0; Gaps 0;
QY 119 VROCXYHGNGQSYRGTSTVTGRTQCSWSSMPHWHQRTPEPNYAGLTENYCRNPADT 178
Db 1 VDDCYHGDGOSYRGTSTVTGRTQCSWSSMPHWHQRTPEPNYAGLTENYCRNPADT 60
QY 179 GPWCFTTDPSPRWEYCNLFQCSSETSGVLEPTVPVPSMEARSAAPTETPVVR 238
Db 61 GPWCFTTDPSPRWEYCNLFQCSSETSGVLEPTVPVPSMEARSAAPTETPVVR 120
QY 239 ATVTGTQCEWAAQEPHRSHTFIPGTNKGAGLEKNCYCRNPDDINGPWCYTMNPKLF 298
Db 121 ATVTGTQCEWAAQEPHRSHTFIPGTNKGAGLEKNCYCRNPDDINGPWCYTMNPKLF 180
QY 299 YCDIPLCASS 307

Db		181 YCDVQCAA 189	!!!: !!:
		AAAB01919 standard; Protein; 192 AA.	
ID	AAAB01919		
AC	AAAB01919;		
XX			
DT	18-SEP-2000 (first entry)		
XX			
DE	Human plasminogen kringle 4-5 (Val355-Phe546).		
XX			
KW	plasminogen; human; kringle domain; endothelial cell proliferation;		
KW	angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;		
KW	antipsoriatic; antiinflammatory; anticulcer; antirheumatic; antiarthritic;		
KW	antiangiogenic; cancer; tumour; autoimmune disease.		
OS	Homo sapiens.		
XX			
PN	US6057122-A.		
XX			
PD	02-MAY-2000.		
XX			
PF	05-MAY-1997; 97US-0851350.		
XX			
PR	03-MAY-1996; 96US-0643219.		
PR	03-APR-1997; 97US-0832087.		
XX			
PA	(ABBO) ABBOTT LAB.		
XX			
PI	Davidson DJ;		
XX			
DR	WPI; 2000-349573/30.		
XX			
PT	Preparation of Kringle five peptide fragment for treating various disorders such as angiogenic, ocular, skin diseases and cancer,		
PT	involves mixing mammalian plasminogen and elastase followed by incubation and isolation -		
PS	Example 17; Page -; 48pp; English.		
XX			
CC	The invention relates to a method of preparing plasminogen kringle 5 peptide fragments. The method comprises mixing mammalian plasminogen and elastase in the ratio 1:100-1:300, followed by incubating and isolating the fragment. The kringle 5 peptides are inhibitors of angiogenesis and endothelial cell proliferation and migration. The peptides are useful for treating angiogenic diseases, primary and metastatic solid tumours and carcinomas of various organs such as breast, genital tract, endocrine glands, skin, tumours of the brain and eyes and solid tumours arising from haematopoietic malignancies such as leukaemias and lymphomas. They are also used for the prophylaxis of various autoimmune diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases (e.g., psoriasis), blood vessel diseases (e.g. haemangiomias, Osler-Webber Syndrome), diseases caused by excessive or abnormal stimulation of endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases which have angiogenesis as a pathologic consequence (e.g., cat scratch disease and ulcers). The peptides are also useful as a birth control agent which inhibits ovulation and establishment of the placenta. Sequences AAB01906-B01919 represent fragments of human plasminogen used in an exemplification of the invention.		
CC	Note: This sequence is not shown in the specification, but is derived from the full length human plasminogen sequence (AAB01887) shown in figure 1.		
XX			
SQ	Sequence 192 AA;		
	Query Watch 51.1%; Score 913; DB 21; Length 192;		
	Best Local Similarity 79.4%; Pred. No. 1.3e-51;		
	Matches 150; Conservative 17; Mismatches 22; Indels 0; Gaps 0;		
YY	119 VROCYNHGNGSQYGTFTSTTGKTCQSWSMTPHRRHQKTPENIPNAGTWTMYCNRNPADK 178		

Query Match 41.7%; Score 744; DB 20; Length 280;
 Best Local Similarity 42.7%; Pred. No. 1.5e-40;
 Matches 131; Conservative 40; Mismatches 100; Indels 36; Gaps 7;

6 QDCYHGDGSRVGISSVTVTGRTCSQSSMIPHHQRTPENYPNAGLTENYCRNPDGSKQ 65
 4 EECMHCSGENDYDKISKTMGLSCQAWDSQSPHAGYIPSKFNKLNKKNYCRNPDREL 63
 66 PWCYTTDPCVRWEYCNLTQCSSETESVLETPVTVVPSMEAHSEAAPTQTPVVRQCYHG 125
 64 PWCFTTDPNKNWELCDIPRCT-----TPPFS-----SGPT-----YQCLKG 99

126 NGOSTRGFTSVTGTGRTCSQSSMIPHHQRTPENYPNAGLTENYCRNPDADTGPWCFTT 185
 100 TGENYRGNAVTVSGHTCQHSQAQTHETHTNTPENFPCKNLNENYCRNPDGKRAPWCHTT 159
 186 DPSIRWEYCNLTGRTSDTEGTGTVVAPPVTVQVPSLGPFS-----EODCMFGNGYGRGKATT 241
 160 NSQVRWEYCKIPSCDS-----SPVSTEQIAPTAPPETLPVWODCYHGDGOSYRGTSSTT 213
 242 VTGTPCQEWAAQOEHRHSTFIPGTNKNWAGLEKKNYCRNPDGDDINGPWCYTMNPKLFYCD 301
 214 TTGKKCQSSMTPHRRQK-TPENYPNAGLTENYCRNPDAD-KGPWCFTTDPVSRWEYCN 271
 302 IPLCASS 308
 272 LKRCST 278

RESULT 5
 AAY02103
 ID AAY02103 standard; Protein; 297 AA.
 AC AAY02103;
 DF 16-JUL-1999 (first entry)
 XX A multifunctional protein of the invention.

Angiostatin; endostatin; interferon; thrombospondin;
 interferon-inducible protein; platelet factor 4; anti-angiogenic;
 anti-tumor; multifunctional protein; angiogenic-mediated disease;
 cancer; diabetic retinopathy; macular degeneration; arthritis;
 tumor cell production.

Synthetic.
 OS Homo sapiens.
 XX W09916889-A1.
 XX 08-APR-1999.
 XX 30-SEP-1998; 98WO-US20464.
 XX 01-OCT-1997; 97US-0060609.
 XX (SEAR) SEARLE & CO G D.
 XX Polanowski MA, Caparon MH, Casperson GF, Gregory SA;
 PI Klein BK, McKearn JP;
 XX WPI; 1999-255098/21.
 XX New multifunctional proteins useful for treating angiogenic-mediated diseases
 XX Claim 5; Page 99-100; 121pp; English.
 XX The specification describes multifunctional proteins which comprise combinations of angiostatin, endostatin, interferon, thrombospondin, interferon-inducible protein and platelet factor 4, and have anti-angiogenic and/or anti-tumor activity. The multifunctional protein may exhibit useful properties such as having similar or greater

CC biological activity when compared to a single factor or by having improved half-life or decreased adverse side effects, or a combination of these properties. The proteins can be used for treating an angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular degeneration, or arthritis. They can also be used for inhibiting the production of tumor cells (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor growth. The present sequence represents a multifunctional protein of the invention.

XX Sequence 297 AA;
 SQ Query Match 41.7%; Score 744; DB 20; Length 297;
 Best Local Similarity 42.7%; Pred. No. 1.6e-40;
 Matches 131; Conservative 40; Mismatches 100; Indels 36; Gaps 7;

6 QDCYHGDGSRVGISSVTVTGRTCSQSSMIPHHQRTPENYPNAGLTENYCRNPDGSKQ 65
 21 EECMHCSGENDYDKISKTMGLSCQAWDSQSPHAGYIPSKFNKLNKKNYCRNPDREL 80
 66 PWCYTTDPCVRWEYCNLTQCSSETESVLETPVTVVPSMEAHSEAAPTQTPVVRQCYHG 125
 81 PWCFTTDPNKNWELCDIPRCT-----TPPFS-----YQCLKG 116
 126 NGOSYRGFTSVTGTGRTCSQSSMTPHRRQRTPENYPNAGLTENYCRNPDADTGPWCFTT 185
 117 TGENYRGNAVTVSGHTCQHSQAQTHETHTNTPENFPCKNLNENYCRNPDGKRAPWCHTT 176
 186 DPSIRWEYCNLTGRTSDTEGTGTVVAPPVTVQVPSLGPFS-----EODCMFGNGYGRGKATT 241
 177 NSQVRWEYCKIPSCDS-----SPVSTEQIAPTAPPETLPVWODCYHGDGOSYRGTSSTT 230
 242 VTGTPCQEWAAQOEHRHSTFIPGTNKNWAGLEKKNYCRNPDGDDINGPWCYTMNPKLFYCD 301
 231 TTGKKCQSSMTPHRRQK-TPENYPNAGLTENYCRNPDAD-KGPWCFTTDPVSRWEYCN 288
 302 IPLCASS 308
 289 LKRCST 295

RESULT 6
 AAW07570
 ID AAW07570 standard; protein; 250 AA.
 AC AAW07570;
 DT 22-JUN-1997 (first entry)
 XX Bovine kringle 1-3.
 XX Angiostatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 KW macular degeneration; diabetic retinopathy.
 XX Bos taurus.
 XX W09635774-A2.
 XX 14-NOV-1996.
 XX 26-APR-1996; 96WO-US05856.
 XX 08-NAR-1996; 96US-0612788.
 XX 26-APR-1995; 95US-0425743.
 XX 22-FEB-1996; 96US-0605598.
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 XX Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
 XX WPI; 1996-518662/51.

PT Use of angiotensin fragments or aggregates - for inhibiting
 PT endothelial cell proliferation and treating angiogenesis-mediated
 PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX
 XX
 PS Claim 4; Page 128; 203pp; English.
 XX
 CC The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotensin fragment, a combination of angiotensin fragments, or
 CC aggregate angiotensin. The fragment is preferably derived from murine,
 CC human, rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotensin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, bovine kringle 1-3, is a specific angiotensin
 CC fragment which can be used in the invention, and represents amino
 CC acids 6-255 of bovine angiotensin.
 XX
 SQ Sequence 250 AA;
 Query Match 40.1%; Score 716; DB 17; Length 250;
 Best Local Similarity 42.1%; Pred. No. 8.6e-39;
 Matches 126; Conservative 43; Mismatches 80; Indels 50; Gaps 7;
 QY 8 CYHGDGRYRGISSTVTGTCTQSSSMIPHWORTPENYPNAGLTPENYCRNPDSGKQ-P 66
 Db 1 CTGTGNGQYRGTTAETKSGVTCTQKWSATSPHVPKFSPEKPLAGLEENYCRNPDSG 60
 QY 67 WCYTDDPCVWEYCNLTQCSSETSGVLETPVTVPVPSMEAHSEAPTEQTPVVRQCYHGN 126
 Db 61 WCYTDDPCVWEYCNLTQCSSETSGVLETPVTVPVPSMEAHSEAPTEQTPVVRQCYHGN 87
 QY 127 GQSYRGTSSTVTGTCTQSSSMIPHWORTPENYPNAGLTPENYCRNPDSGKQ-P 186
 Db 88 GENYEGKIAKTSQSGDQANDSQSPHAGYTPSKFPNKNLKNYCRNPDSGKQ-P 147
 QY 187 PSIRWEYCNLTQCSSETSGVLETPVTVPVPSMEAHSEAPTEQTPVVRQCYHGN 246
 Db 148 POKRWEFCIDPRCT-----TPP-----PSSGP--KYQCLKGTGKNGYGVAVTESGHT 193
 QY 247 CQWAAQEPHRSHTFIPGTNKGLEKNYCRNPDSGKQ-P 305
 Db 194 CQWSEQEPHKNR--TPENFCKNLEENYCRNPDSG--KAPWCITTNSEVREYETCTPSC 250

RESULT 7
 AA07569
 ID AA07569 standard; protein; 250 AA.
 XX
 AC AA07569;
 XX
 DT 22-JUN-1997 (first entry)
 XX
 DE Porcine kringle 1-3.
 XX
 KW angiotensin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 KW macular degeneration; diabetic retinopathy.
 XX
 OS Sus scrofa.
 XX
 PN W09635774-A2.
 XX
 PD 14-NOV-1996.
 XX
 PF 26-APR-1996; 96WO-US05856.
 XX
 PR 08-MAR-1996; 96US-0612788.
 PR 26-APR-1995; 95US-0429743.

PR 22-FEB-1996; 96US-0605598.
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 XX
 PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
 XX WPI; 1996-518662/51.
 XX
 DR Use of angiotensin fragments or aggregates - for inhibiting
 XX endothelial cell proliferation and treating angiogenesis-mediated
 XX diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX
 PS Claim 4; Page 126-127; 203pp; English.
 XX
 CC The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotensin fragment, a combination of angiotensin fragments, or
 CC aggregate angiotensin. The fragment is preferably derived from murine,
 CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotensin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, porcine kringle 1-3, is a specific angiotensin
 CC fragment which can be used in the invention, and represents amino
 CC acids 6-255 of porcine angiotensin.
 XX
 SQ Sequence 250 AA;
 Query Match 40.0%; Score 715; DB 17; Length 250;
 Best Local Similarity 42.5%; Pred. No. 1e-38;
 Matches 127; Conservative 41; Mismatches 81; Indels 50; Gaps 7;
 QY 8 CYHGDGRYRGISSTVTGTCTQSSSMIPHWORTPENYPNAGLTPENYCRNPDSG-KOP 66
 Db 1 CTGTGNGQYRGTTAETKSGVTCTQKWSATSPHVPKFSPEKPLAGLEENYCRNPDSG 60
 QY 67 WCYTDDPCVWEYCNLTQCSSETSGVLETPVTVPVPSMEAHSEAPTEQTPVVRQCYHGN 126
 Db 61 WCYTDDPCVWEYCNLTQCSSETSGVLETPVTVPVPSMEAHSEAPTEQTPVVRQCYHGN 87
 QY 127 GQSYRGTSSTVTGTCTQSSSMIPHWORTPENYPNAGLTPENYCRNPDSGKQ-P 186
 Db 88 GENYEGKIAKTSQSGDQANDSQSPHAGYTPSKFPNKNLKNYCRNPDSGKQ-P 147
 QY 187 PSIRWEYCNLTQCSSETSGVLETPVTVPVPSMEAHSEAPTEQTPVVRQCYHGN 246
 Db 148 POKRWEFCIDPRCT-----TPP-----PSSGP--KYQCLKGTGKNGYGVAVTESGHT 193
 QY 247 CQWAAQEPHRSHTFIPGTNKGLEKNYCRNPDSGKQ-P 305
 Db 194 CQWSEQEPHKNR--TPENFCKNLEENYCRNPDSG--KAPWCITTNSEVREYETCTPSC 250

RESULT 8
 AA02109
 ID AA02109 standard; protein; 268 AA.
 XX
 AC AA02109;
 XX
 DT 16-JUL-1999 (first entry)
 XX
 DE A multifunctional protein of the invention.
 XX
 KW Angiotensin; endostatin; interferon; thrombospondin;
 KW interferon-inducible protein; platelet factor 4; anti-angiogenic;
 KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
 KW cancer; diabetic retinopathy; macular degeneration; arthritis;
 KW tumor cell production.

X Purifying recombinant angiostatin, involves applying fermented broth
T containing angiostatin to expanded bed cation exchange column, anion
T exchange column, hydroxyapatite column, hydrophobic column and a
T membrane
X
S Example 1; Page 22; 49pp; English.
X
C The present sequence representing Angiostatin protein is given in an
C invention providing a method for recombinant production, recovery and
C purification of Angiostatin protein. Purification of recombinant
C Angiostatin comprises applying crude fermentation broth containing the
C protein to an expanded bed cation exchange column, eluting it, and
C applying the eluate to anion exchange column, repeating the process of
C eluting and applying, to hydroxyapatite column, hydrophobic column and
C membrane, in order, and collecting fluid passing through the membrane.
C Angiostatin is useful for treating angiogenesis mediated diseases,
C including solid tumours, leukaemia, tumour metastases, benign tumours,
C rheumatoid arthritis, psoriasis, ocular angiogenic diseases, Osler-Webber
C syndrome, myocardial angiogenesis, plaque neovascularisation,
C telangiectasia, haemophilic joints, angiofibroma and wound granulation.
C As a centrifugation technique is not employed in the process, damage to
C the cells with concomitant release of undesirable biological materials
C such as cytochromes, pigments, enzymes, chemicals and other undesirable
C cellular constituents and debris, is prevented. Large scale recovery and
C purification of proteins is greater than that obtained from prior art
C methods. Active Angiostatin can be stored in buffers for extended periods
C of time, in vials or other containers, either in solution which may be
C liquid or frozen, or lyophilised.
X
Q Sequence 260 AA;

Query Match 39.8%; Score 711; DB 22; Length 260;
Best Local Similarity 42.3%; Pred. No. 1.9e-38;
Matches 129; Conservative 40; Mismatches 86; Indels 50; Gaps 7;
Y 5 VQCYHGDSYRGISSTVTGRCQSWSSMTPHWHQRTPEYFNAGLTENYCRNPDGSK 64
b 1 LSECKTGNGKYGTPMSKTKNGITCQKWSSTSPHPRFSPATHFSEGLEENYCRNPDNDP 62
Y 65 Q-PWCYTTPDCVRYEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCQY 123
b 63 QGPWCYTTPDCVRYEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCQY 89
Y 124 HNGQSYRGTFSTVTGRCQSWSSMTPHWHQRTPEYFNAGLTENYCRNPDADGPKCF 183
b 90 HCSGNYDGRISKTSGLEQAWDSQSPHAGYIPSKFNNKLNKKNYCRNPDRELPRWCF 149
Y 184 TTPDSIRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCQY 243
b 150 TTPDNKRWELCDIPRCT-----TPP-----PSSGPTYQ--CLKGTGNYRGNVAVTVS 195
Y 244 GTPCQWAAQEPHRSHTFIPGTNKGWGLEKNYCRNPDGNGPCWGTMPNPKLFYCDIP 303
b 196 GHPCQWAAQEPHRSHTFIPGTNKGWGLEKNYCRNPDGNGPCWGTMPNPKLFYCDIP 253
Y 304 LCASS 308
b 254 SCDS 258

RESULT 14

AY53869

D AAY53869 standard; protein; 254 AA.

C AAY53869;

X AAY53869;

T 13-MAR-2000 (first entry)

X Amino acid sequence of human greenstatin protein.

E Human; plasminogen; angiostatin; greenstatin; thrombolytic factor;

W angiogenesis inhibitory protein; proliferation; angiogenesis; cancer;

N

KW vascular endothelial cell; ophthalmic disease; glaucoma;
KW diabetic retinopathy; arthritis; psoriasis.
XX Homo sapiens.
OS WO9961464-A1.
XX 02-DEC-1999.
XX 28-MAY-1999; 99WO-KR00263.
XX 28-MAY-1998; 98KR-0019335.
XX 27-MAY-1999; 99KR-0019144.
PA (GREG) KOREA GREEN CROSS CORP.
PI You WK, So SH, Ahn BC, Lee H, Jung S, Kim Y, Lee JH, Hong Y;
PI Joe YA, Chang S;
XX WPI; 2000-086703/07.
XX Purifying angiogenesis inhibitors produced as recombinant proteins in
PT Escherichia coli, useful as anticancer agents and for treating ocular
PT diseases
PS Claim 2; Page 47-48; 55pp; English.
XX The present sequence represents the human greenstatin protein. It
CC is derived from the plasminogen protien, and comprises amino acids
CC 101-354. Angiostatin is also derived from plasminogen, and comprises
CC amino acids 99-467. Angiostatin and greenstatin are used as
CC thrombolytic factors and angiogenesis inhibitory proteins. Angiostatin
CC contains the kringle 1-4 region of plasminogen, and greenstatin contains
CC the kringle 1-3 region of plasminogen. As both proteins contain a high
CC number of disulphide bonds, they are difficult to purify. The
CC specification describes a method for the purification of such
CC angiogenesis inhibitory proteins. The method comprises solubilising
CC the proteins, produced as inclusion bodies in Escherichia coli and
CC refolding the solubilised fraction in buffer containing urea and
CC glutathione. The angiogenesis inhibitory proteins specifically inhibit
CC proliferation of vascular endothelial cells, but not that of
CC non-endothelial cancers or normal cells. The angiogenesis inhibitory
CC proteins are used to suppress angiogenesis, specifically for treating
CC cancers (e.g. of lung, skin and brain) or ophthalmic diseases (e.g.
CC glaucoma and diabetic retinopathy), but also arthritis and psoriasis.
XX Sequence 254 AA;
SQ
Query Match 39.6%; Score 707; DB 21; Length 254;
Best Local Similarity 42.4%; Pred. No. 3.3e-38;
Matches 128; Conservative 40; Mismatches 84; Indels 50; Gaps 7;
QY 7 DCYHGDGRISYRGISSTVTGRCQSWSSMTPHWHQRTPEYFNAGLTENYCRNPDGSKQ- 65
DB 2 ECKTGNGKYGTPMSKTKNGITCQKWSSTSPHPRFSPATHFSEGLEENYCRNPDNDPQ 61
QY 66 PWCYTTPDCVRYEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCQY 125
DB 62 PWCYTTPDCVRYEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCQY 88
QY 126 NGQSYRGTFSTVTGRCQSWSSMTPHWHQRTPEYFNAGLTENYCRNPDADGPKCF 185
DB 89 SGENYDGRISKTSGLEQAWDSQSPHAGYIPSKFNNKLNKKNYCRNPDRELPRWCF 148
QY 186 DPISIRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCQY 245
DB 149 DPKNRWELCDIPRCT-----TPP-----PSSGPTYQ--CLKGTGNYRGNVAVTVSGH 194
QY 246 PQEWAAQEPHRSHTFIPGTNKGWGLEKNYCRNPDGNGPCWGTMPNPKLFYCDIP 305
DB 195 TQWAAQEPHRSHTFIPGTNKGWGLEKNYCRNPDGNGPCWGTMPNPKLFYCDIP 252
QY 306 AS 307

Db	88	GENYDGTSTKMSGLEQAWDSQSPHAGHYIPSKFNKNUKKNVCNPDGEPWCMETTD	147
QY	187	PSTRWEYCNLTRGSDTEGVVAPPVTVQVSLGSPSEQDCMFGNGKGYRGKATVTGTP	246
Db	148	PNKRMELDIPRCH-----TTP-----PSSGPYQ--CLKGTGENYGDVATVVSQHT	193
QY	247	QEWAAQSPHRHSTFIPTGNKWAGLEKNYCRNDEGSDINGPWCYTMNPKLFYDCDILPC	305
Db	194	CHWGAQSPHPHNR--PENFPCKNLDENYCRNPDGE--KAPWCYTTSOVKWEYKPIESC	250

Search completed: September 29, 2003, 08:03:03
Job time : 45 secs

b) 253 DS 254

RESULT 15
AAW07568

D AAW07568 standard; protein; 250 AA.
C AC
X AAW07568;
Y 22-JUN-1997 (first entry)
Z X Rhesus kringle 1-3.
DE X
E angiotensin; plasminogen; kringle; angiogenesis; cancer; arthritis;
F macular degeneration; diabetic retinopathy.
G W
H Rhesus.
I NS WO9635774-A2.
J N 14-NOV-1996.
D Y
F 26-APR-1996; 96WO-US05856.
T R 08-MAR-1996; 96US-0612788.
R 26-APR-1995; 95US-0429743.
R 22-FEB-1996; 96US-060598.
A (CHIL-) CHILDRENS MEDICAL CENT.
I Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
T WFJ; 1996-518662/51.
R
X Use of angiotensin fragments or aggregates - for inhibiting endothelial cell proliferation and treating angiogenesis-mediated diseases, e.g. cancer, arthritis or diabetic retinopathy
T T
T Claim 4; Page 125-126; 203pp; English.
S
X The invention relates to new methods and compositions for inhibiting endothelial cell proliferation, using as active component an angiotensin fragment, a combination of angiotensin fragments, or aggregate angiotensin. The fragment is preferably derived from murine, human, rhesus, porcine or bovine plasminogen and is a kringle 1, kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle 1-4 or kringle 1-4BKLS protein. The aggregate angiotensin has a Mol. wt. of 45-65 kD and is derived from a plasminogen fragment beginning at approximately amino acid number 98 of murine, human, rhesus, porcine or bovine plasminogen. The active component can be used for treating angiogenesis-mediated diseases such as cancer, arthritis, macular degeneration and diabetic retinopathy. It can also be used to develop antibodies for use in diagnosis, detection and therapy.
C The present sequence, Rhesus kringle 1-3, is a specific angiotensin fragment which can be used in the invention, and represents amino acids 6-255 of Rhesus angiotensin.

X Z Sequence 250 AA;

Query Match 39.4%; Score 704; DB 17; Length 250;
Best Local Similarity 42.5%; Pred. No. 5.le-38;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

Y 8 CYHGDSRYRGISSTVTGTGRCSWSMTPHHQRTPEYNAGLTENYCRNPD-SKOP 66
 | : : : : : | : : : : : | : : : : : | : : : : : |
c 1 CKTGNGKNRGMTSKRGITGCQSWESTSPHRPTFSFPATPSGLEENYCRNPNDQGQF 60

Y 67 WCYTDPDCVRWCNTTCQSETESGVLFTPTVPFPSNEAHSEAPTSQTIVPQQCYHEN 126
 ||||| | : : : : : | : : : : : | : : : : : | : : : : : |
b 61 WCYTDPEDRFCDIPCE-----ECMHCS 87

127 GSYSRGISTVTGTGRCSWSMTPHHQRTPEYNPNGLTMNYCRNPADTGPWFCTTD 186

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:01:08 ; Search time 29 seconds
(without alignments)
449.370 Million cell updates/sec

Title: US-10-088-548-2

Perfect score: 1786

Sequence: 1 KSPVQDCYHGDSYRGIS.....YTMNPKLFYCDIPLCASS 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 282628

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCBUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	716	40.1	250	2	US-08-612-788-33
2	716	40.1	250	3	US-09-066-028-33
3	716	40.1	250	4	US-09-335-325-33
4	715	40.0	250	2	US-08-612-788-32
5	715	40.0	250	3	US-09-066-028-32
6	715	40.0	250	4	US-09-335-325-32
7	704	39.4	250	2	US-08-612-788-31
8	704	39.4	250	3	US-09-066-028-31
9	704	39.4	250	4	US-09-335-325-31
10	703	39.4	250	2	US-08-612-788-30
11	703	39.4	250	3	US-09-066-028-30
12	703	39.4	250	4	US-09-335-325-30
13	700	39.2	250	2	US-08-612-788-29
14	700	39.2	250	3	US-09-066-028-29
15	700	39.2	250	4	US-09-335-325-29
16	585	32.8	210	3	US-08-985-526-21
17	538	30.1	168	2	US-08-612-788-24
18	538	30.1	168	3	US-09-066-028-24
19	538	30.1	168	4	US-09-335-325-24
20	531	29.7	168	2	US-08-612-788-27
21	531	29.7	168	3	US-09-066-028-27
22	531	29.7	168	4	US-09-335-325-27
23	524	29.3	130	4	US-08-240-839-4
24	519	29.1	91	3	US-09-234-553-1
25	516	28.9	120	1	US-08-211-747-7
26	515	28.8	168	2	US-08-612-788-26
27	515	28.8	168	3	US-09-066-028-26

28	515	28.8	168	4	US-09-335-325-26
29	514	28.8	168	2	US-08-612-788-28
30	514	28.8	168	3	US-09-066-028-28
31	514	28.8	168	4	US-09-335-325-28
32	511	28.6	168	2	US-08-612-788-25
33	511	28.6	168	3	US-09-066-028-25
34	511	28.6	168	4	US-09-335-325-25
35	503	28.2	160	2	US-08-612-788-38
36	503	28.2	160	3	US-09-066-028-38
37	503	28.2	160	4	US-09-335-325-38
38	499	27.9	160	2	US-08-612-788-37
39	499	27.9	160	3	US-09-066-028-37
40	499	27.9	160	4	US-09-335-325-37
41	490	27.4	160	2	US-08-612-788-36
42	490	27.4	160	3	US-09-066-028-36
43	490	27.4	160	4	US-09-335-325-36
44	489	27.4	160	2	US-08-612-788-35
45	489	27.4	160	3	US-09-066-028-35

ALIGNMENTS

RESULT 1

US-08-612-788-33
; Sequence 33, Application US/08612788
; Patent No. 5837682

GENERAL INFORMATION:

APPLICANT: Folkman, M. Judah
APPLICANT: O'Reilly, Micheal
APPLICANT: Cao, Yihai
APPLICANT: Sim, B. Kim Lee
TITLE OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE:
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 250 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Bovine
IMMEDIATE SOURCE:
CLONE: KI-3
US-08-612-788-33

US-08-612-788-33

```
Query Match 40.1%; Score 716; DB 2; Length 250;
Best Local Similarity 42.1%; Pred. No. 1.8e-57;
Matches 126; Conservative 43; Mismatches 80; Indels 50; Gaps 7;

iy 8 CYHGDGRYRIGTSITVTGRTCSWSSMIPRHQRTPENYNAGLTENYCRNPDGSKQ-P 66
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | | | |
ib 1 CKTGNGQYRGTAAETKSGVTCQKWSATSPHVPKFSPEKPLAGLEENYCRNPDNGP 60
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | | |
iy 67 WYTTDPCVWEYCNLTQCSSESVLETPVVPVPSMEAHSEAAPEQTVPVROCYHGN 126
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
ib 61 WYTTDPCVWEYCNLTQCSSESVLETPVVPVPSMEAHSEAAPEQTVPVROCYHGN 126
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
iy 127 GQSYRGTSFTVTGRTCSWSSMIPRHQRTPENYNAGLTENYCRNPDGSKQ-P 186
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
ib 88 GENYEGKIATMSGRDQAWDSQSPHAGYIPSKFPNKLKMYCNRPDGEPRWCFTTD 147
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
iy 187 PSIRWEYCNLTQCSSESVLETPVVPVPSMEAHSEAAPEQTVPVROCYHGN 246
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
ib 148 PQRWFECDIPRCT-----TPP-----PSSGP--KYQCLKGTGKNGYGVAVTESGHT 193
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
iy 247 CQEWAAQEPHRSFTIPGTNKGLEKNYCRNPDGDPINGPWCYTMNPKLFYCDIPLC 305
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
ib 194 CQWSEQTPHKNR--TPENFPCKNLEENYCRNPDGE--KAPWCYTTNSENVEYCIIPSC 250
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |

RESULT 2
US-09-066-028-33
; Sequence 33, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
```

```
; ORGANISM: Bovine
; IMMEDIATE SOURCE:
; CLONE: K1-3
US-09-066-028-33

Query Match 40.1%; Score 716; DB 3; Length 250;
Best Local Similarity 42.1%; Pred. No. 1.8e-57;
Matches 126; Conservative 43; Mismatches 80; Indels 50; Gaps 7;

QY 8 CYHGDGRYRIGTSITVTGRTCSWSSMIPRHQRTPENYNAGLTENYCRNPDGSKQ-P 66
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
Db 1 CKTGNGQYRGTAAETKSGVTCQKWSATSPHVPKFSPEKPLAGLEENYCRNPDNGP 60
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
QY 67 WYTTDPCVWEYCNLTQCSSESVLETPVVPVPSMEAHSEAAPEQTVPVROCYHGN 126
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
Db 61 WYTTDPCVWEYCNLTQCSSESVLETPVVPVPSMEAHSEAAPEQTVPVROCYHGN 126
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
QY 127 GQSYRGTSFTVTGRTCSWSSMIPRHQRTPENYNAGLTENYCRNPDGSKQ-P 186
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
Db 88 GENYEGKIATMSGRDQAWDSQSPHAGYIPSKFPNKLKMYCNRPDGEPRWCFTTD 147
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
QY 187 PSIRWEYCNLTQCSSESVLETPVVPVPSMEAHSEAAPEQTVPVROCYHGN 246
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
Db 148 PQRWFECDIPRCT-----TPP-----PSSGP--KYQCLKGTGKNGYGVAVTESGHT 193
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
QY 247 CQEWAAQEPHRSFTIPGTNKGLEKNYCRNPDGDPINGPWCYTMNPKLFYCDIPLC 305
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |
Db 194 CQWSEQTPHKNR--TPENFPCKNLEENYCRNPDGE--KAPWCYTTNSENVEYCIIPSC 250
   | | | | | : : : | | | | | : : | | | | | | | | | | | | | |

RESULT 3
US-09-335-325-33
; Sequence 33, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
```

GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
APPLICANT: O'Reilly, Michael
APPLICANT: Cao, Yihai
APPLICANT: Sim, B. Kim Lee
TITLE OF INVENTION: Angiotatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids

```

RESULT 5
US-09-066-028-32
; Sequence 32, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiotstatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:

```

```

/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 32:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Porcine
/ IMMEDIATE SOURCE:
/ CLONE: K1-3
/ SEQUENCE DESCRIPTION: SEQ ID NO: 32:
/
US-09-335-325-32
/
Query Match 40.0%; Score 715; DB 4; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.2e-57;
Matches 127; Conservative 41; Mismatches 81; Indels 50; Gaps 7;

QY 8 CYHGDRSTRGISSTVTGRTGCSWSMTPHHQRTPTENYPNAGLTENICRPDGS-QQP 66
DB 1 CKYGNKNGRTSTKTSYGICQWSYSSPHIKYSPKFFPLAGLENYICRPNDDEKGP 60
QY 67 WCYTPDPCVRYCNLTGCGTESGVLETPTVVPVPSMEAHSEAAPEQTTPVRCYHGN 126
DB 61 WCYTPDPRFYCDIPECD-----ECMHCS 87
QY 127 GQSYRGFTFTVTGRTGCSWSMTPHHQRTPTENYPNDGLTMYNCRPNADTGPWCFTTD 186
DB 88 GEHYEGKISKTMSTGTCQWSGSPAHGYLPSKEPNKLNKMYNCRPNDEGPRPWCFTTD 147
QY 187 PSTRWYCNLTGSDTEGTVVADPTVIQVPSLGGPSEODCMFGNGKGYRGKATTVTGP 246
DB 148 PNRKWFCDIPACT-----TPGTYQ---CLKGGENYRTVSVFASGHT 193
QY 247 CQWAAQAEPRHSTETPGTNKNWAGLEKNCYCRPDGDIINGPWCYTMNPKLFYCDIPLC 305
DB 194 CQWQAQSPHKHNR-TPENPPCKNLENYICRPDGE-TAPWCYTTDSEVWDYCKIPSC 250

RESULT 7
US-08-612-788-31
/ Sequence 31, Application US/08612788
/ Patent No. 5837682
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:

```

```

/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE:
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: K1-3
/ US-08-612-788-31

```

```

Query Match 39.4%; Score 704; DB 2; Length 250;
Best local similarity 42.5%; Pred. No. 2.2e-56;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY 8 CYHGDGRSGYRGSSIVTGRTCQSWSMIPHHQRTPENYPNAGLTENYCRNPDG-KGOP 66
Db 1 CKTGNGKNGYGTMSKTRGICQKWSSTSPHRFTFSPATHPSEGLEENYCRNPDGQGP 60
QY 67 WCYTTPDCVWVEYCNLTQCSSETESGVLETPVVPVPSMEAHSEAAPTQTPTVVRQCYHGN 126
Db 61 WCYTTPDEERFDYCDIPECD-----ECMHCS 87
QY 127 GQSYRGFTSTVYTGRTQCSWSMTPHRHQRTPENYPNDGLTMYNCRNPDADTGPWCFTTD 186
Db 88 GENTDGRKISKTMSGLEQAWDSQSPHAGYIPSKFPNKLKNYCRNPDGEPWPWCFTTD 147
QY 187 PSIRWEYCNLTGCSDEGTGVVAPPTVIQVPSLGPSPQDCMFGNGKGYRGKATTVTGT 246
Db 148 PNKEWELCDIPRCT-----TPP-----PSSGPTYQ--CLAGTGENYRGDVAVTVSGHT 193
QY 247 CQEWAAQEPHRHSTPIPTGNKAGLEKNYCRNPDGDIINGFWCYTNPNPKLFDYCDIPLC 305
Db 194 CHGWSAQTPHTNR-TPENFFCKNLNENYCRNPDGE-KAPWCYTTTNSQVRWEYCKIPSC 250

```

```

RESULT 8
US-09-066-028-31
/ Sequence 31, Application US/09066028
/ Patent No. 6024688
/ GENERAL INFORMATION:
/ APPLICANT: Polkman, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiotensin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/066,028
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/612,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: K1-3
/ US-09-066-028-31

Query Match 39.4%; Score 704; DB 3; Length 250;
Best local similarity 42.5%; Pred. No. 2.2e-56;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY 8 CYHGDGRSGYRGSSIVTGRTCQSWSMIPHHQRTPENYPNAGLTENYCRNPDG-KGOP 66
Db 1 CKTGNGKNGYGTMSKTRGICQKWSSTSPHRFTFSPATHPSEGLEENYCRNPDGQGP 60
QY 67 WCYTTPDCVWVEYCNLTQCSSETESGVLETPVVPVPSMEAHSEAAPTQTPTVVRQCYHGN 126
Db 61 WCYTTPDEERFDYCDIPECD-----ECMHCS 87
QY 127 GQSYRGFTSTVYTGRTQCSWSMTPHRHQRTPENYPNDGLTMYNCRNPDADTGPWCFTTD 186
Db 88 GENTDGRKISKTMSGLEQAWDSQSPHAGYIPSKFPNKLKNYCRNPDGEPWPWCFTTD 147
QY 187 PSIRWEYCNLTGCSDEGTGVVAPPTVIQVPSLGPSPQDCMFGNGKGYRGKATTVTGT 246
Db 148 PNKEWELCDIPRCT-----TPP-----PSSGPTYQ--CLAGTGENYRGDVAVTVSGHT 193
QY 247 CQEWAAQEPHRHSTPIPTGNKAGLEKNYCRNPDGDIINGFWCYTNPNPKLFDYCDIPLC 305
Db 194 CHGWSAQTPHTNR-TPENFFCKNLNENYCRNPDGE-KAPWCYTTTNSQVRWEYCKIPSC 250

```

```

RESULT 9
US-09-335-325-31
/ Sequence 31, Application US/09335325
/ Patent No. 6521439
/ GENERAL INFORMATION:
/ APPLICANT: Polkman, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiotensin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.

```

```

;
;
; ZIP: 30303-1769
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
;
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rhesus monkey
; IMMEDIATE SOURCE:
; CLONE: K1-3
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-335-325-31

```

```

Query Match 39.4%; Score 704; DB 4; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.2e-56;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY 8 CYHGRSIRGSSITVTGRTCSWSSMTPHQRHQTPTNPYCNAGLTENYCRNPDGSGKQ-P 66
Db 1 CKTGKNGYRTGMSRTGTCOKWSSTSPHRTSPATHPSEGLEENYCRNPDNDGQGP 60
QY 67 WCYTDPQVRWEYCNLTQCSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCYHGN 126
Db 61 WCYTDPQVRWEYCNLTQCSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCYHGN 87
QY 127 GQSYRGTFSTVTGRTCSWSSMTPHQRHQTPTNPYCNAGLTENYCRNPDADTGPWCFTTD 186
Db 88 GENTYDGIKSTMSGLEQANDSQSPAHGYIPSKFPNKLKKNYCRNPDGPRPWCFTTD 147
QY 187 PSIRWEYCNLTQCSDEGTGVVAPPTVIVPSLGPPSEODCMFGNGKGYRGKATTVTGP 246
Db 148 PNKRWELCDIPRCT-----TPP-----PSSGPTYQ--CLKGTGENYRGVAVTVSGHT 193
QY 247 CQWAAQEPHRSSTPIPTGNKAGLEKNYCRNPDGDINGPWCYTMNPKLFYCDIPLC 305
Db 194 CHGWSAQTPHTNHR--TPENFPCKNLNENYCRNPDGE--KAPWCYTTNSQVWEYCKIPSC 250

RESULT 10
US-08-612-788-30
; Sequence 30, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45

```

```

;
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
;
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K1-3
;
; US-08-612-788-30

```

```

Query Match 39.4%; Score 703; DB 2; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.7e-56;
Matches 127; Conservative 39; Mismatches 83; Indels 50; Gaps 7;

QY 8 CYHGRSIRGSSITVTGRTCSWSSMTPHQRHQTPTNPYCNAGLTENYCRNPDGSGKQ-P 66
Db 1 CKTGKNGYRTGMSRTGTCOKWSSTSPHRTSPATHPSEGLEENYCRNPDNDGQGP 60
QY 67 WCYTDPQVRWEYCNLTQCSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCYHGN 126
Db 61 WCYTDPQVRWEYCNLTQCSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCYHGN 87
QY 127 GQSYRGTFSTVTGRTCSWSSMTPHQRHQTPTNPYCNAGLTENYCRNPDADTGPWCFTTD 186
Db 88 GENTYDGIKSTMSGLEQANDSQSPAHGYIPSKFPNKLKKNYCRNPDRLRPWCFTTD 147
QY 187 PSIRWEYCNLTQCSDEGTGVVAPPTVIVPSLGPPSEODCMFGNGKGYRGKATTVTGP 246
Db 148 PNKRWELCDIPRCT-----TPP-----PSSGPTYQ--CLKGTGENYRGVAVTVSGHT 193
QY 247 CQWAAQEPHRSSTPIPTGNKAGLEKNYCRNPDGDINGPWCYTMNPKLFYCDIPLC 305
Db 194 CHGWSAQTPHTNHR--TPENFPCKNLNENYCRNPDGE--KAPWCYTTNSQVWEYCKIPSC 250

RESULT 11
US-09-066-028-30
; Sequence 30, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee

```



```
;; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
;; NUMBER OF SEQUENCES: 45
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Jones & Askew
;; STREET: 191 Peachtree Street, 37th Floor
;; CITY: Atlanta
;; STATE: Georgia
;; COUNTRY: U.S.
;; ZIP: 30303-1769
;; COMPUTER READABLE FORM:
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/066,028
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/612,788
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren, William L.
;; REGISTRATION NUMBER: 36,714
;; REFERENCE/DOCKET NUMBER: 05213-0126
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 404-818-3700
;; TELEFAX: 404-818-3799
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 250 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; IMMEDIATE SOURCE:
;; CLONE: K1-3
;;
US-09-066-028-30
Query Match 39.4%; Score 703; DB 3; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.7e-56;
Matches 127; Conservative 39; Mismatches 83; Indels 50; Gaps 7;

QY 8 CYHGDGRSYRIGSTTTVTGTCQSWSSMIPHWHQRTPEYFNAGLTENYCRNPDSGKQ-P 66
DB 1 CKTGNGKYNRGTMSKTKNGITCQKWSSTSPHPRFSPATHPSGLENYCRNPDPQGP 60
QY 67 WCYTDDPCVWEYCNLTQCSSETESGVLETPVVPVPSMEAHSEAAPTETPTVVRQCHGN 126
DB 61 WCYTDDPKRYDYCDILECEE-----ECMHCS 87
QY 127 GOSYRGFTSTVTGTCQSWSSMTPHWHQRTPEYFNAGLTENYCRNPDPADTGPWCFTTD 186
DB 88 GENYDGKISKYMSGLSCQANDSQSPHAGYIPSKFPKNLKNKYNCRNPDRRLRWCFTTD 147
QY 187 PSIRWEYCNLTFRCSDTETGTVVAPPTVIOVPSLGPPSDQCMFGNGKYGKKAFTVGTGP 246
DB 148 PNKRWELCDIPRCT-----TPP-----PSSGPTYQ--CLKGTGENYRGNVAVTSHT 193
QY 247 CQWAAQEPHRSHTFIPGTHKAGLEKNYCRNPDPDINGPWCYTMNPKLFYDCIDPLC 305
DB 194 CQWAAQEPHRSHTFIPGTHKAGLEKNYCRNPDPDINGPWCYTMNPKLFYDCIDPLC 305
; Sequence 30, Application US/09335325
; Patent No. 6521439
```

RESULT 12

US-09-335-325-30

; Sequence 30, Application US/09335325

; Patent No. 6521439

```
;; GENERAL INFORMATION:
;; APPLICANT: Folkman, M. Judah
;; O'Reilly, Micheal
;; Cao, Yihai
;; Sim, B. Kim Lee
;; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
;; NUMBER OF SEQUENCES: 45
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Jones & Askew
;; STREET: 191 Peachtree Street, 37th Floor
;; CITY: Atlanta
;; STATE: Georgia
;; COUNTRY: U.S.
;; ZIP: 30303-1769
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/335,325
;; FILING DATE: 17-Jun-1999
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/612,788
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren, William L.
;; REGISTRATION NUMBER: 36,714
;; REFERENCE/DOCKET NUMBER: 05213-0126
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 404-818-3700
;; TELEFAX: 404-818-3799
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 250 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: <Unknown>
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; IMMEDIATE SOURCE:
;; CLONE: K1-3
;; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-335-325-30
Query Match 39.4%; Score 703; DB 4; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.7e-56;
Matches 127; Conservative 39; Mismatches 83; Indels 50; Gaps 7;

QY 8 CYHGDGRSYRIGSTTTVTGTCQSWSSMIPHWHQRTPEYFNAGLTENYCRNPDSGKQ-P 66
DB 1 CKTGNGKYNRGTMSKTKNGITCQKWSSTSPHPRFSPATHPSGLENYCRNPDPQGP 60
QY 67 WCYTDDPCVWEYCNLTQCSSETESGVLETPVVPVPSMEAHSEAAPTETPTVVRQCHGN 126
DB 61 WCYTDDPKRYDYCDILECEE-----ECMHCS 87
QY 127 GOSYRGFTSTVTGTCQSWSSMTPHWHQRTPEYFNAGLTENYCRNPDPADTGPWCFTTD 186
DB 88 GENYDGKISKYMSGLSCQANDSQSPHAGYIPSKFPKNLKNKYNCRNPDRRLRWCFTTD 147
QY 187 PSIRWEYCNLTFRCSDTETGTVVAPPTVIOVPSLGPPSDQCMFGNGKYGKKAFTVGTGP 246
DB 148 PNKRWELCDIPRCT-----TPP-----PSSGPTYQ--CLKGTGENYRGNVAVTSHT 193
QY 247 CQWAAQEPHRSHTFIPGTHKAGLEKNYCRNPDPDINGPWCYTMNPKLFYDCIDPLC 305
DB 194 CQWAAQEPHRSHTFIPGTHKAGLEKNYCRNPDPDINGPWCYTMNPKLFYDCIDPLC 305
; Sequence 30, Application US/09335325
; Patent No. 6521439
```

```
RESULT 13
US-08-612-788-29
; Sequence 29, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K1-3
US-08-612-788-29

Query Match 39.28; Score 700; DB 2; Length 250;
Best Local Similarity 41.18; Pred. No. 5e-56;
Matches 123; Conservative 37; Mismatches 89; Indels 50; Gaps 6;

QY 8 CYHGGRSIRGISTVTGRTQCSWSSMIPHHQRTPEPNYPNAGLTENYCRNPDGSKQ-P 66
Db 1 CKTGIGNGYRGTMSTKSGVACQKKGATFPHPVNYSPSTHPNGLNENYCRNPDNDGCP 60
QY 67 WCYTDPDPCVRWEYCNLQCSSETSGVLETPVTPVPPSMEASAPTEQTPTVVRQCYHGN 126
Db 61 WCYTDPDKRYDYNICEPEE-----ECMYCS 87
QY 127 GQSYRGTFTVTGRTQCSWSSMTPRHQRTPEPNYPNDGLTMNCRNPADTGPWCFTTD 186
Db 88 GEYEGKISKTMISGLDQANDSQSPHAGYIPAKFPKSKNLMKANYCHNPDPGPRWCFTTD 147
QY 187 PSIRWEYCNLRCSDTGTGTVVAPPTVIQPSLGPFPSCQCFNGKGYRGKATTVTGTG 246
Db 148 PTKRWEYCDIPRC-----TTPPPP-----PSPTYQLKGRGENYGTGTVTSVSKT 193
QY 247 CQWRAQEPHRSIFPGTKNAGLEKNYCRNPNPDGDLNGPWCYTMTNPKLFYCDIPLC 305

RESULT 14
US-09-066-028-29
; Sequence 29, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K1-3
US-09-066-028-29

Query Match 39.28; Score 700; DB 3; Length 250;
Best Local Similarity 41.18; Pred. No. 5e-56;
Matches 123; Conservative 37; Mismatches 89; Indels 50; Gaps 6;

QY 8 CYHGGRSIRGISTVTGRTQCSWSSMIPHHQRTPEPNYPNAGLTENYCRNPDGSKQ-P 66
Db 1 CKTGIGNGYRGTMSTKSGVACQKKGATFPHPVNYSPSTHPNGLNENYCRNPDNDGCP 60
QY 67 WCYTDPDPCVRWEYCNLQCSSETSGVLETPVTPVPPSMEASAPTEQTPTVVRQCYHGN 126
Db 61 WCYTDPDKRYDYNICEPEE-----ECMYCS 87
QY 127 GQSYRGTFTVTGRTQCSWSSMTPRHQRTPEPNYPNDGLTMNCRNPADTGPWCFTTD 186
Db 88 GEYEGKISKTMISGLDQANDSQSPHAGYIPAKFPKSKNLMKANYCHNPDPGPRWCFTTD 147
```


Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	716	40.1	250	9	US-09-761-120-33	Sequence 33, Appl
2	716	40.1	250	10	US-09-335-325-33	Sequence 33, Appl
3	716	40.1	250	15	US-10-131-241-33	Sequence 33, Appl
4	715	40.0	250	9	US-09-761-120-32	Sequence 32, Appl
5	715	40.0	250	10	US-09-335-325-32	Sequence 32, Appl
6	715	40.0	250	15	US-10-131-241-32	Sequence 32, Appl
7	711	39.8	260	15	US-10-131-241-61	Sequence 61, Appl
8	704	39.4	250	9	US-09-761-120-31	Sequence 31, Appl
9	704	39.4	250	10	US-09-335-325-31	Sequence 31, Appl
10	704	39.4	250	15	US-10-131-241-31	Sequence 31, Appl
11	703	39.4	250	9	US-09-761-120-30	Sequence 30, Appl
12	703	39.4	250	10	US-09-335-325-30	Sequence 30, Appl
13	703	39.4	250	15	US-10-131-241-30	Sequence 30, Appl
14	700	39.2	250	9	US-09-761-120-29	Sequence 29, Appl
15	700	39.2	250	10	US-09-335-325-29	Sequence 29, Appl


```

; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rhesus monkey
; IMMEDIATE SOURCE:
; CLONE: KI-3
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-335-325-31

```

```

Query Match      39.4%; Score 704; DB 10; Length 250;
Best Local Similarity 42.5%; Fred. No. 8.6e-54;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY      8 CYHGDGRSYRGISSTVTGRTCSQSSMIPHWHQRTPENYPNAGLTENYCRNPDG-KQGP 66
Db      1 CKTGKNKRYGTMSTRTGTCQKWSSTSPHRTFTSPATHPSEGLEENYCRNPDNDGQGP 60

QY      67 WCYTDPGVWVEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPTQTPVVRQCYHGN 126
Db      61 WCYTDPGVWVEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPTQTPVVRQCYHGN 126

QY      127 GQSYRGFTSTVTGRTCSQSSMIPHWHQRTPENYPNAGLTENYCRNPDADTGPWCFTTD 186
Db      88 GENDGKISKTMSGLECCQAWDSQSPHAGHYIPSKFFKNLKNYCRNPDGPRPWCFTTD 147

QY      187 PSIRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPTQTPVVRQCYHGN 246
Db      148 PNKRWELCDIPRCT-----TPP-----PSSGPTQY--CLAGTGENYRGDVAVTVSGHT 193

QY      247 COEWAQEPHRSHTPIPTGNKAGLEKNYCRNPDGDPINGPWCYTNNPKLFYDCIDPLC 305
Db      194 CHGWSAQTPHTNHR--TPENFCKNLNENYCRNPDGE--KAPWCYTINSQVRWEYCKIPSC 250

RESULT 11
US-09-761-120-30
; Sequence 30, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Polkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 30
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 1-3

```

```

; Sequence 31, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Rhesus monkey
; OTHER INFORMATION: US-10-131-241-31

```

```

Query Match      39.4%; Score 704; DB 15; Length 250;
Best Local Similarity 42.5%; Fred. No. 8.6e-54;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY      8 CYHGDGRSYRGISSTVTGRTCSQSSMIPHWHQRTPENYPNAGLTENYCRNPDG-KQGP 66
Db      1 CKTGKNKRYGTMSTRTGTCQKWSSTSPHRTFTSPATHPSEGLEENYCRNPDNDGQGP 60

QY      67 WCYTDPGVWVEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPTQTPVVRQCYHGN 126
Db      61 WCYTDPGVWVEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPTQTPVVRQCYHGN 126

QY      127 GQSYRGFTSTVTGRTCSQSSMIPHWHQRTPENYPNAGLTENYCRNPDADTGPWCFTTD 186
Db      88 GENDGKISKTMSGLECCQAWDSQSPHAGHYIPSKFFKNLKNYCRNPDGPRPWCFTTD 147

QY      187 PSIRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPTQTPVVRQCYHGN 246
Db      148 PNKRWELCDIPRCT-----TPP-----PSSGPTQY--CLAGTGENYRGDVAVTVSGHT 193

QY      247 COEWAQEPHRSHTPIPTGNKAGLEKNYCRNPDGDPINGPWCYTNNPKLFYDCIDPLC 305
Db      194 CHGWSAQTPHTNHR--TPENFCKNLNENYCRNPDGE--KAPWCYTINSQVRWEYCKIPSC 250

```

```

RESULT 11
US-09-761-120-30
; Sequence 30, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Polkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 30
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 1-3

```


Db 88 GENTDGRKISKTSLGLEQAWDSQSPHAGYIPKFPNKLKNKYNCRNPDRPLPWCFTTD 147
QY 187 PSIRWEYCNLTRCSDEGTGVAPPVIVQVPSLGPSPQDCMFGNGKGYRGKATVGTGP 246
Db 148 PNKEWELCDIPRCT-----TPP-----PSSGPTYQ--CLAGTGENYRGNAVTVSGHT 193
QY 247 COEWAAQEPHSHSTTFPGINKWAGLEKNYCRNPDPDINGPWCYTMNPKRLFDYCDIPLC 305
Db 194 CQWSEQTPHHRN--TPENFPCKNLNENYCRNPDPGE--RAPWCHTINSQVRWEYCKIPSC 250

RESULT 14
US-09-761-120-29
; Sequence 29, Application US/097611120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Kringle 1-3
US-09-761-120-29

Query Match 39.2%; Score 700; DB 9; Length 250;
Best Local Similarity 41.1%; Pred. No. 1.9e-53;
Matches 123; Conservative 37; Mismatches 89; Indels 50; Gaps 6;

QY 8 CYHGDGRSYRGISSTVTGRTQSQSSMTPHWHQRTPTENYPNAGLTENYCRNPDSGKO-P 66
Db 1 CKTGIGNGYRGTSRKTSGVACQKWGATFPHPVNPSPSTHPNGLNENYCRNPDEQGP 60
QY 67 WCYTTPDCVRWEYCNLTQCSSETESGVLETPVVPVPSMEAHSEAAPEQTVPVRCYHGN 126
Db 61 WCYTTPDKRKYDYNICEPEE-----ECMVC 87
QY 127 GOSYRGFTSTVTGRTQSQSSMTPHWHQRTPTENYPNAGLTENYCRNPDPADTGPWCFTTD 186
Db 88 GEKYEGKISKTMSGLDCQAWDSQSPHAGYIPAKFPKLNKMYCHNPDPGEPRPWCFTTD 147
QY 187 PSIRWEYCNLTRCSDEGTGVAPPVIVQVPSLGPSPQDCMFGNGKGYRGKATVGTGP 246
Db 148 PTKRWEYCDIPRC-----ATPPPP-----PSPYQCLKGRGENYRGVTSVTSVSGKT 193
QY 247 COEWAAQEPHSHSTTFPGINKWAGLEKNYCRNPDPDINGPWCYTMNPKRLFDYCDIPLC 305
Db 194 CQWSEQTPHHRN--TPENFPCKNLNENYCRNPDPGE--TAPWCYTTPDSQLRWEYCEIPSC 250

RESULT 15
US-09-335-325-29
; Sequence 29, Application US/09335325
; Patent No. US2002016471A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiotatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K1-3
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-335-325-29

Query Match 39.2%; Score 700; DB 10; Length 250;
Best Local Similarity 41.1%; Pred. No. 1.9e-53;
Matches 123; Conservative 37; Mismatches 89; Indels 50; Gaps 6;

QY 8 CYHGDGRSYRGISSTVTGRTQSQSSMTPHWHQRTPTENYPNAGLTENYCRNPDSGKO-P 66
Db 1 CKTGIGNGYRGTSRKTSGVACQKWGATFPHPVNPSPSTHPNGLNENYCRNPDEQGP 60
QY 67 WCYTTPDCVRWEYCNLTQCSSETESGVLETPVVPVPSMEAHSEAAPEQTVPVRCYHGN 126
Db 61 WCYTTPDKRKYDYNICEPEE-----ECMVC 87
QY 127 GOSYRGFTSTVTGRTQSQSSMTPHWHQRTPTENYPNAGLTENYCRNPDPADTGPWCFTTD 186
Db 88 GEKYEGKISKTMSGLDCQAWDSQSPHAGYIPAKFPKLNKMYCHNPDPGEPRPWCFTTD 147
QY 187 PSIRWEYCNLTRCSDEGTGVAPPVIVQVPSLGPSPQDCMFGNGKGYRGKATVGTGP 246
Db 148 PTKRWEYCDIPRC-----ATPPPP-----PSPYQCLKGRGENYRGVTSVTSVSGKT 193
QY 247 COEWAAQEPHSHSTTFPGINKWAGLEKNYCRNPDPDINGPWCYTMNPKRLFDYCDIPLC 305
Db 194 CQWSEQTPHHRN--TPENFPCKNLNENYCRNPDPGE--TAPWCYTTPDSQLRWEYCEIPSC 250

Search completed: September 29, 2003, 08:18:17
Job time : 349 secs

A:Reference number: A61545; MUID:89005015; PMID:3168975

A:Accession: E61545
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-120 <SCH>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: hydrolase; serine proteinase
F:37-114/Domain: kringle homology <KR4>

Query Match 23.1%; Score 413; DB 2; Length 120;
Best Local Similarity 74.7%; Pred. No. 7.le-24;
Matches 65; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 119 VRQCCHNGSGKGFSTVTGRTCSQSSMTPHRHQRTPENPDGLTMYCRNPDA DT 178
 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 34 VQECHNGSGRGTSSTITIGRCQSWSMTPHREKTPEHFPEAGLTMYCRNPDA K 93
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 179 GPWCFTDPSIRWEYNLTRGSDEGT 205
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 94 SPWCYTDPVSRWFCNLRKLCPAS 120
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3

C61545

plasmin (EC 3.4.21.7) precursor - goat (fragments)
N:Alternate names: plasminogen
C:Species: Capra aegagrus hircus (domestic goat)
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C:Accession: C61545
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: C61545
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-123 <SCH>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: hydrolase; serine proteinase
F:41-118/Domain: kringle homology <KR4>

Query Match 22.8%; Score 408; DB 2; Length 123;
Best Local Similarity 73.9%; Pred. No. 1.7e-23;
Matches 68; Conservative 9; Mismatches 13; Indels 2; Gaps 1;

QY 110 AAAPTEQPVVROCIHGNSGYRGFTSVTGRTCQSWSMTPHRHQRTPEINPDGLTN 169
 || : : : | | | | | | | | | | | | | | | | | | | | | | |
Db 31 AKACEEE--AQDCIHGNSGYRGTSVTVGRCQSWSSKTPHRRHKPTPEYNAGLTKN 88
 || : : : | | | | | | | | | | | | | | | | | | | | | | |

QY 170 YCRNPDA DTGPWCFTDPSIRWEYNLTRCS D 201
 ||||| | | | | | | | | | | | | | | | | | | | | | | |
Db 89 YCRNPDA DKSPWCYTDPVRWRWFCKLKGE 120
 ||||| | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4

A60140

plasmin (EC 3.4.21.7) precursor - chicken (fragment)
N:Alternate names: plasminogen
C:Species: Gallus gallus (chicken)
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C:Accession: A60140
R:Gyenes, M.; Paththy, L.
Biochim. Biophys. Acta 832, 326-330, 1985
A:title: The kringle 4 domain of chicken plasminogen.
A:Reference number: A60140; MUID:86077796; PMID:4074753
A:Accession: A60140
A:Molecule type: protein
A:Residues: 1-89 <GYE>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; plasma; serine proteinase
F:6-83/Domain: kringle homology <KRG>
F:6-83,21-66,55-78/Disulfide bonds: #status predicted
F:39/Binding site: carbohydrate (Asn) (covalent) #status experimental

```

Query Match          19.8%; Score 354; DB 2; Length 89;
Best Local Similarity 67.9%; Pred. No. 1.2e-19;
Matches 55; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 121 QCVHNGQSVRGFTSTVTGTGTCQSSSTPHRHORTPENYNDGLTMMAYCRNPADATGCP 180
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 ECVQNGVSVRGFTASTTITGKKCAQNMSSPHRHKNTESHFPNADLRQNYCRNPADAKSP 64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 WCFTTDPSTRWECYCNLCRCD 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 WCYTTDSVRYWCYCNLCRCD 85
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C:date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human en
A:Reference number: I38098; MCID:90192128; PMID:1969145
A:Accession: I38098
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <STE>
A:Cross-references: EMBL:X13097; NID:g35282; PIDN:CAA31489.1; PID:g35283
C:Comment: For the main splice form, see PIR:UKHUT. This form probably does not have
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:14-32/Domain: propeptide #status predicted <PRO>
F:33-291/Product: t-plasminogen activator, inactive endothelial splice form #status p
F:41-78/Domain: fibronectin type I repeat homology <IFA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KRL1>
F:215-291/Domain: kringle homology #status atypical <K32>
F:41-71, 69-78, 86-97, 91-108, 110-119, 127-208, 148-190, 179-203/Disulfide bonds: #status p

Query Match          13.8%; Score 246.5; DB 2; Length 291;
Best Local Similarity 26.2%; Pred. No. 3.7e-11;
Matches 80; Conservative 34; Mismatches 110; Indels 81; Gaps 14;

```

```

14 RSYRGI---SSTVTGRTCQWSNAPHHQRTPTNYPNAGLTENYCRNPDSKGPWCYT 70
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
35 RSYOVCROEKQMTIYOQHOSWLPVLR--SNRVEYCWCSNGRAQCHSPVVKSCSPFCFN 93
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
71 TDPVCWRVEYN--LTQCSETSGVLEPTVWPVPSMEAHSEAAFTPTVWPQCYHNGQ 128
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
94 GGTCCQALFISDFVCCQFEGAG-----KCEIDTRAT-----CTEDOGI 133
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
129 SYRGTFSTVTIGRTCQWSNMTPHRQ---RTPENYPNDGLTM-----NYCRNPDAITGP 180
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
134 SYRGTWSTAESGAECTNWNNSALAQNAYSRRP----DAIRLGLGNHNYCRNPDRDSKP 188
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
181 WCFTTDP--SIRWEYCNLRCSDBTEGTWVAPPVIOVPSLGPSPQDCMFGNKGVRGKA 239
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
189 WCYVFRAGKYSFEFCSTPACSE-----GNSDCTFGNGSAYRGTFS 228
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
240 TTVTGRPPCQEW-----AAQEPHRHSTIPGNTKNWAGLEK--NYCRNPDDGJINGWC 288
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
229 LTESGASCLFWNSMLIGKYTAQNPSAQ-----LGLGKHNYCRT--GRSYSSPAT 278
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
289 YTMNP 293
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```


C;Keywords: glycoprotein; transmembrane protein
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.8%; Score 86.5; DB 2; Length 297;
Best Local Similarity 23.0%; Pred. No. 26;
Matches 57; Conservative 26; Mismatches 88; Indels 77; Gaps 13;
QY 66 PWCYTDPCVRWEYCNLQCSETSGVLETPVVPVPSMEAHSEAAPEQTVPVRCYHG 125
Db 88 PTVLQNPQLGHSFNL---SETTSQPAATPAIT-TPSAES-----TPGSTVVK----- 132
QY 126 NGQSYRGTSITVTGTCQSSMTPHRHQRTPEYNDGL---TMNYCRNPADATGPMC 182
Db 133 -----TKNTTIT---QIQP-SKPTTKQHQKPPKPNKHFFVFNFPSCISNNPTC 182
QY 183 F-----TTDPSIRWEYCNLTRCSDEGTGVVAPTVTVQVSLGPPSE 223
Db 183 WAJCKRIPKPKKTKTKPKTKI-----TTKDLKPKTPKPEVLTT-----KPT 232
QY 224 QDCMFGNGKGYRGKATP-VTGTPCQEWAAQEPHRSTFIPGCTKNWAGLEKNYCRNPDDG 282
Db 233 KPINTTKNIRITLLTNTTNGNP--EYTSQKETLHST-----SPGNG 273
QY 283 INGPWCYT 290
Db 274 PPSQVYT 281

RESULT 10
AB3430
transposase BMEI1424 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AB3430
R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanova,
; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AB3430
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-272 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL52605.1; PID:g17983424; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1424
A;Map position: 1

Query Match 4.8%; Score 85.5; DB 2; Length 272;
Best Local Similarity 21.8%; Pred. No. 28;
Matches 62; Conservative 29; Mismatches 77; Indels 117; Gaps 18;
QY 30 QSSSMTPHWHQRTPEYVNPAGLNEY-----CRNPSGKQPCWYTTDCV- 75
Db 10 QQWAVIAP-----LLPNQPGAHRTDDRVSIGIHLRSGCRWQDC---PACIGPPTVY 62
QY 76 ----RWF-----YCNLTQCSSETSGVLETPVVPVPSMEAHSEAA----- 111
Db 63 NRFTWAKGIWRLEALVQPTDRDTHMIDSTTA-----KAHPPAAGGKGDAEAI 116
QY 112 -----PTEQPPVVRQCYHGNGQSYGT-----FSTTV-----TGRT 142
Db 117 SRGGRSTKIAVVDSC--GRPVALRITPGQGDAPITVILLDVSQEVVHSEIETGVT 174
QY 143 QSSSMTPHWHQRTPEYVNPAGLNEY-----MNYCRNPADATG---PW-----CFTDPS 188
Db 175 TPQSFRIAPNEH---PQECPTDALASRNGRSGFRQADGAGCTPLMGLAENCFLSDRT 231
QY 189 IR--WEYCNLTRCSDEGTGVVAPTVTVQVSLGP-----PSEQ 224
Db 232 VODIWP-----GGDDGLSVAPH--OPPSNGFGHCRITDTPSQ 267

RESULT 11

AB5231
hypothetical protein AY4920350 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: AB5231
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: AB5001; MUID:20083488; PMID:10617198
A;Accession: AB5231
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-241 <STO>
A;Cross-references: GB:NC_001268; NID:g7768830; PIDN:CAB79035.1; GSPDB:GN00140
C;Genetics:
A;Gene: AY4920350
A;Map position: 4

Query Match 4.7%; Score 84.5; DB 2; Length 241;
Best Local Similarity 21.8%; Pred. No. 29;
Matches 42; Conservative 16; Mismatches 70; Indels 65; Gaps 9;
QY 95 TPTVTPVPSMEAHSEAAPEQTVPVRCYHGNGQSRGTFTTGTGTCQSSMTPHRH 154
Db 11 TPTVTPPGF-----ITDEEQTLNHHYGASGSKWK-----TLNRRRLQNWGMV-HEK 59
QY 155 QRTPEYVNP-----NDGLTMNYCRNPADATGPGWCTTDPDSIRWEY 193
Db 60 GLVPQELFPWLTKTAIEHSSGLFPAINHVLNEY---HPDQGIMP----- 104
QY 194 CNLTRCSDEGTGVVAPTVTVQVPSLGGPPSQD---CMFNGKGYRGKATVTVTPCOE 249
Db 105 -----HQDGPAYFP--VVALISLGSFVVMDFTPHLRLSRGDSGYISKDQ-----SPCAE 150
QY 250 WAAQEPHRHSTFI 262
Db 151 SCAPERDSFVLL 163

RESULT 12

T24703
hypothetical protein T08G3.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T24703
R;Lloyd, C.
submitted to the EMBL Data Library, December 1996
A;Reference number: Z19926
A;Accession: T24703
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-304 <WIL>
A;Cross-references: EMBL:283238; PIDN:CAB05796.1; GSPDB:GN00023; CESP:T08G3.6
A;Experimental source: clone T08G3
C;Genetics:
A;Gene: CESP:T08G3.6
A;Map position: 5
A;Introns: 86/2

Query Match 4.7%; Score 84.5; DB 2; Length 304;
Best Local Similarity 19.1%; Pred. No. 37;
Matches 53; Conservative 33; Mismatches 108; Indels 83; Gaps 14;
QY 87 ETESGVLETVVVPV-----PMEAHSEAAPEQ--TP-----VTRQC 122
Db 26 EINFSEIFKTRIVPLMSVQAGCGVSGAPRRNALQNAQKICTPGSACRQISLIVRVC 85
QY 123 YHNGQYRGTFSTTFTV-----GRFC-----QSWSSMTPHRHQRTPEY-YPND 164
Db 86 ---XGDTYRILCNKTCVPAQRTCCIPVYVPMINGSQCGCFPPDPDRSESEFCPCPKG 142


```
QY 165 GLTMN---YCRNPDAFGPW-----CFTDPSIRWEYCNLTFRCSDTGTVVAP 209
Db 143 GFWSWWSAYFR--DGEKKAWSTRRLCTEEAGCCTDPTGTTIETSTACFCRLKLVDSLV 200
QY 210 PTVIQVPSLGGP--SEQDCMFGNGKGYGKATTVTGTPCOEAAQPHRHSTFI-----PG 264
Db 201 KTNLGTYYLTPVYNDEDCASQVM---ESQKSAVGECPCNEYISWYIYNTWNAVRYKPG 256
QY 265 TKNWAGLKNYCRNPDGDINGPCWYTMNPKLFYCD 301
Db 257 RSDYGVKISNCE-----AAHEKKIFLYCD 281

RESULT 13
JOI208
attachment protein - human respiratory syncytial virus (strain RS6256)
N:Alternate names: G protein
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
C:Accession: JOI208
R:Canine, P.A.; Matthews, D.A.; Pringle, C.R.
J. Gen. Virol. 72, 2091-2096, 1991
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A
A:Reference number: JOI204; MUID:91374005; PMID:1895054
A:Accession: JOI208
A:Molecule type: mRNA
A:Residues: 1-297 <CAN>
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill-
children and adults.
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
C:Keywords: glycoprotein; transmembrane protein
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.7%; Score 83.5; DB 2; Length 297;
Best Local Similarity 22.6%; Pred. No. 43;
Matches 56; Conservative 26; Mismatches 89; Indels 77; Gaps 13;

QY 66 PNCYTDDCVRWECNLTQCSFESGVLETPVVPVPSMEAHSEAAFTQTTPVRCQYHG 125
Db 88 PYYLQNQLGISFNL---SETTS-----QPTTTPAPTTPS---ABSTFQSTTVK----- 132

QY 126 NGOSVGTFTVTGRCQSWSSMTPHRHQRTPENYPNDGL---TMNYCRNPDAFGPW 182
Db 133 -----TKNITT--QIOP-SKPTTKQKQKPPKPNNDPHEVFNFPVPCSTCSNNPTC 182

QY 183 F-----TTDPSIRWEYCNLTFRCSDTGTVVAPTVIQVPSLGGPSE 223
Db 183 WAICKRIPNKKRPGKNTTKTKPTIK-----TTKDLKPQTTKPEVLTT-----KPT 232

QY 224 QDCMFGNGKGYGKATTVTGTPCOEAAQPHRHSTFIPGTNKNWAGLKNYCRNPDGD 282
Db 233 KPTINTRTNRTILLTNTGNP--EYTSQKETHLST-----SPEGN 273

QY 283 INGFWCYT 290
Db 274 PPSQVYT 281

RESULT 14
T33698
hypotheical protein F49F1.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 08-Sep-2000
C:Accession: T33698
R:Miller, N.; Wamsley, P.
submitted to the EMBL Data Library, October 1998
A:Description: the sequence of C. elegans cosmid F49F1.
A:Reference number: Z21389
A:Accession: T33698
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-240 <MIL>
A:Cross-references: EMBL:AF100656; PIDN:AA68946.1; GSPDB:GN00022; CESP:F49F1.7
```

```
A:Experimental source: strain Bristol N2; clone F49F1
C:Genetics:
A:Gene: CESP:F49F1.7
A:Map position: 4
A:Introns: 42/2; 79/2; 116/2; 209/2
C:Superfamily: Caenorhabditis elegans hypothetical protein F49F1.6

Query Match 4.6%; Score 82; DB 2; Length 240;
Best Local Similarity 21.1%; Pred. No. 44;
Matches 44; Conservative 17; Mismatches 74; Indels 74; Gaps 9;

QY 28 TCQSWSSMIPHWHQRT-----PENYPNAGLTENYCRNPDGSKQPCWYCTTDPVCRWEYCNL 82
Db 57 TCSDINSCEKFEHMCVMPYRQCPKS-----CGICDSPATWCM-----HWPECPK 105

QY 83 TC-CSEFTE-----SGVLETPVVPVPSMEAHSEAAFTQT 116
Db 106 YQVPCSETERIKCPRSCDAGNPSTKEASVTITNGHLQIGTTRRLQATQITKSTETTT 165

QY 117 PVRQCYHNGQSRYGTFSTVTGTCQSWSSMTPHRHQRTPENYPNDGLTMNYCRNDA 176
Db 166 ETK-----AIKTKTETTKATTTTPSTVTKTTKP-----KPECT 204

QY 177 DTGPPW-----FTTD-----PSIRWEYC 194
Db 205 DSSPNCITWAKNGFCTNTFYPPKREYC 233

RESULT 15
TI9913
hypotheical protein C43F9.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: TI9913
R:Mortimore, B.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19195
A:Accession: TI9913
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-248 <WIL>
A:Cross-references: EMBL:Z82262; PIDN:CAB05152.1; GSPDB:GN00022; CESP:C43F9.5
A:Experimental source: clone C43F9
C:Genetics:
A:Gene: CESP:C43F9.5
A:Introns: 47/1; 78/3; 118/1; 199/3

Query Match 4.5%; Score 81; DB 2; Length 248;
Best Local Similarity 23.8%; Pred. No. 55;
Matches 59; Conservative 19; Mismatches 114; Indels 56; Gaps 15;

QY 13 GRSYRGIS--TTVGTCTQSWSSMIP--HHQRTPENYPNAGLTENYCRNPDGSKQFW 67
Db 20 GLTELGIKSKRQTHICGTYPNQFISMPCEYWTSTQVFN-----PYTCANGGRKIGV 72

QY 68 CYTTDDPCVW---EYCNLTQCSFESGVLETPVVPVPSMEAHSEAAFTQTTPVRCQYH 124
Db 73 CYTNQCTPYAANSVC-LNNCCCTNPVVTTRAPITTT-----TFSVALAYCYN 122

QY 125 GNQSYRGTFST-TVTGTCQSWSSMTPHRHQRTPENYPND--GL-TMNYCRNPDAFGP 180
Db 123 GORTVQRTTSDVCAAGOTCMNGICCT-----TTGNEYTGSCGGLPAISAC-----GTGQ 172

QY 181 WC---FTTDPSTWEYCNLTFRCSDTGTVVAPTVIQVPSLGGPSEQDCMFGNGKGYGKK 238
Db 173 TCGSEFCTSS---NYCECQYGTAG-----LCSNGCPTGYSCNSNTNGY--CC 217

QY 239 ATTVTGTP 246
Db 218 ATCAGRP 225
```

Wed Oct 1 15:58:32 2003

us-10-088-548-2.sep29.rpr

Page 6

Search completed: September 29, 2003, 08:04:08
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:01:08 ; Search time 11 Seconds
(without alignments)
1316.748 Million cell updates/sec

Title: US-10-088-548-2

Perfect score: 1786

Sequence: 1 KSPVQDCYHGDSRGIS.....YTMNPKLFYCDIFLCASS 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 65706

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	620	34.7	169	1 PLMN_RAT	Q01177 rattus norv
2	88.5	5.0	277	1 TNRA_HUMAN	P43489 homo sapien
3	86.5	4.8	297	1 VGLG_HRSV7	P27026 human respi
4	84.5	4.7	221	1 GSHE_MOUSE	P21765 mus musculu
5	83.5	4.7	297	1 VGLG_HRSV6	P27025 human respi
6	79	4.4	221	1 GSHE_RAT	P30710 rattus norv
7	77.5	4.3	221	1 VGLG_MACFA	P28714 macaca fasc
8	75.5	4.2	292	1 VGLG_HRSV8	P23041 human respi
9	75	4.2	127	1 WAP_RABIT	P09412 oryctolagus
10	75	4.2	221	1 GSHE_PIG	O64625 rattus norv
11	74.5	4.2	219	1 GSHE_PIG	O18994 sus scrofa
12	74	4.1	224	1 XP4_XENLA	O00223 xenopus lae
13	74	4.1	271	1 TOLB_BACSU	P42413 bacillus su
14	73.5	4.1	292	1 VGLG_HRSV1	P20896 human respi
15	73	4.1	265	1 KNHL_CANGA	O74884 candida gla
16	73	4.1	300	1 TR6B_HUMAN	O95407 homo sapien
17	72.5	4.1	217	1 SGSS_DROSI	P13729 drosophila
18	72.5	4.1	221	1 GSHE_HUMAN	O75715 homo sapien
19	72.5	4.1	269	1 WBP1_HUMAN	O96627 homo sapien
20	72	4.0	263	1 SGSH_HUMAN	P22352 homo sapien
21	72	4.0	272	1 TNRA_MOUSE	P47741 mus musculu
22	72	4.0	297	1 VGLG_HRSV4	P27023 human respi
23	71.5	4.0	162	1 HXA9_CAVPO	P51783 cavia porce
24	71.5	4.0	221	1 GSHE_CANFA	O46607 canis fami
25	71	4.0	226	1 GSHP_HUMAN	P22352 homo sapien
26	71	4.0	227	1 CTG5_HUMAN	O9br10 homo sapien
27	71	4.0	308	1 HMB1_STRPU	P13545 strongyloce
28	70.5	3.9	169	1 GEO_SIN3	P74250 synecocyst
29	70.5	3.9	261	1 Y612_SYNY3	P72581 synecocyst
30	70	3.9	241	1 WAP8_HUMAN	O81uao4 homo sapien
31	69.5	3.9	283	1 HXA9_FUGRU	O42506 fuqu rubrip
32	69.5	3.9	307	1 HMB1_TRIGR	P09080 tritonuete
33	69	3.9	199	1 EOST_ACTEQ	P81439 actinia equ

34	69	3.9	237	1 IPDE_DICDI	P22549 dictyosteli
35	69	3.9	272	1 ERG_MOUSE	P81270 mus musculu
36	68.5	3.8	278	1 VGLL_HCMV2	Q68668 human cytom
37	68	3.8	275	1 UL11_HCMV2	P16721 human cytom
38	67.5	3.8	226	1 GSHP_RAT	P23764 rattus norv
39	67.5	3.8	278	1 VGLL_HCMV1	Q68667 human cytom
40	67.5	3.8	304	1 YQOB_CREBL	Q09300 caenorhabdi
41	67	3.8	154	1 GP41_BPSP1	O48397 bacterioph
42	67	3.8	238	1 ET3_HUMAN	P14138 homo sapien
43	66.5	3.7	258	1 PRIA_LENED	Q01200 leontinula e
44	66.5	3.7	278	1 VGLL_HCMV4	Q68670 human cytom
45	66.5	3.7	278	1 VGLL_HCMV5	Q68671 human cytom

ALIGNMENTS

RESULT 1					
PLMN_RAT					
ID	PLMN_RAT	STANDARD;	PRT;	169 AA.	
AC	Q01177;				
DT	01-APR-1993 (Rel. 25, Created)				
DT	01-APR-1993 (Rel. 25, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Plasminogen (EC 3.4.21.7) (Fragment).				
GN	PLG.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
KC	TISSUE=Liver;				
KX	MEDLINE=91250378; PubMed=1645711;				
RA	Kanalas J.J.; Makker S.P.;				
RT	*Identification of the rat Heymann nephritis autoantigen (GP330) as a				
RT	receptor site for plasminogen.*;				
RL	J. Biol. Chem. 266:10825-10829(1991).				
CC	-I- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS				
CC	A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING				
CC	EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,				
CC	AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE				
CC	GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN				
CC	ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH				
CC	AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,				
CC	LAMININ AND VON WILLEBRAND FACTOR.				
CC	-I- CATALYTIC ACTIVITY: Preferential cleavage: Lys- -Xaa > Arg- -Xaa;				
CC	higher selectivity than trypsin. Converts fibrin into soluble				
CC	products.				
CC	-I- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN				
CC	ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO				
CC	FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.				
CC	-I- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN				
CC	IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.				
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.				
CC	-I- SIMILARITY: Contains 5 kringle domains.				

This SWISS-PROT entry is copyright. It is produced through a collaboration					
between the Swiss Institute of Bioinformatics and the EMBL outstation -					
the European Bioinformatics Institute. There are no restrictions on its					
use by non-profit institutions as long as its content is in no way					
modified and this statement is not removed. Usage by and for commercial					
entities requires a license agreement (See http://www.isb-sib.ch/announce/					
or send an email to license@isb-sib.ch).					

EMBL:	M62832;	AAA41884.1;	-		
PIR:	A40522;	A40522.			
DR	HSP;	P00747;	IPMK.		
DR	MEROFS;	S01.233;	-		
DR	InterPro:	IPR000001;	Kringle.		
DR	InterPro:	IPR003966;	Prothrombin.		
DR	InterPro:	IPR001254;	Ser.protease_Try.		
DR	Pfam;	PF00051;	kringle; 2.		

```

DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE.1; 1.
DR PROSITE; PS00070; KRINGLE.2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; PARTIAL.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON_TER 169 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;

Query Match 34.7%; Score 620; DB 1; Length 169;
Best Local Similarity 61.0%; Pred. No. 1.9e-42;
Matches 111; Conservative 14; Mismatches 43; Indels 14; Gaps 2;

QY 76 RWECNLTQCSSETSGVLETPVVPVPSNEAHSEAPTEQTPTVVRQCYHGNGOSYRGTFES 135
DQ 1 RWECETPSGSGVSPQSDSSVLB-----EQTPVVECYQGGKSGYRTSS 47
QY 136 TTVGTRCQSNMTPRHQRTPEPNYNDGLTMYNCRNPDAD-TGPGWFTTDPISRWEYC 194
DQ 48 TTTGKCKQSWMTPHSESKPAPNFPSDGLMNYCRNPNDQPGFWCFTTDPISRWEYC 107
QY 195 NLTRCSTEGTVAPTVTVQVPSLGPSPSEODCFMGKGYRGKATTVTGTGPOEWAQAE 254
DQ 108 NLKRCSETGGVAESIAVQVPSAPGTSETDCMTGNGKEYRGTAVTAAGTPOEWAQAE 167
QY 255 PH 256
DQ 168 PH 169

RESULT 2
TNR4_HUMAN STANDARD; PRT; 277 AA.
AC P43489; Q13663;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 4 precursor (OX40L
DE glycoprotein 1 receptor) (TAX-transcriptionally activated
DE glycoprotein 1 receptor) (CD134 antigen).
GN TNFRSF4 OR TXGP1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE-94170844; PubMed=7510240;
RA Latza U., Duerkop H., Schmittger S., Ringeling J., Eitelbach F.,
RA Hummel M., Fonatsch C., Stein H.;
RT "The human OX40 homolog: cDNA structure, expression and chromosomal
RT assignment of the ACT35 antigen.";
RL Eur. J. Immunol. 24:677-683(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-95219871; PubMed=7704935;
RA Baum P.R., Gayle R.B. III, Ramsdell F., Srinivasan S., Sorensen R.A.,
RA Watson M.L., Seldin M.F., Clifford K.N., Grabstein K., Alderson M.R.;
RT "Identification of OX40 ligand and preliminary characterization of
RT its activities on OX40 receptor.";
RL Circ. Shock 44:30-34(1994).
```

```

RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-20318724; PubMed=10861060;
RA Pankow R., Duerkop H., Latza U., Krause H., Kunzendorf U., Pohl T.,
RA Bulfone-Paus S.;
RT "The HVLV-I protein transcriptionally modulates OX40 antigen
RT expression.";
RL J. Immunol. 165:263-270(2000).
RN [4]
RP INTERACTION WITH TRAF1; TRAF2 AND TRAF3.
RX PubMed=9418902;
RA Arch R.H., Thompson C.B.;
RT "4-LBB and OX40 are members of a tumor necrosis factor (TNF)-nerve
RT growth factor receptor subfamily that bind TNF receptor-associated
RT factors and activate nuclear factor kappaB.";
RL Mol. Cell. Biol. 18:558-565(1998).
RN [5]
RP INTERACTION WITH TRAF2 AND TRAF5.
RX PubMed=9488716;
RA Kawamata S., Hori T., Imura A., Takaoi-Kondo A., Uchiyama T.;
RT "Activation of OX40 signal transduction pathways leads to tumor
RT necrosis factor receptor-associated factor (TRAF) 2- and
RT TRAF5-mediated NF-kappaB activation.";
RL J. Biol. Chem. 273:5808-5814(1998).
CC -!- FUNCTION: Receptor for TNFRSF4/OX40L/GP34.
CC -!- SUBUNIT: Interacts with TRAF2, TRAF3 and TRAF5.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 4 TNFR-Cys repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD134 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd134.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75962; CAA53576.1; -.
DR EMBL; S76792; AAB33944.1; ALT_INIT.
DR EMBL; AJ277151; CAB96543.1; -.
DR PIR; I37552; I37552.
DR RSP; 014763; ID06.
DR GENE; HGNC:11918; TNFRSF4.
DR MIM; 600315; -.
DR GO; GO:0005687; C:integral to plasma membrane; TAS.
DR GO; GO:0005027; F:NGF/TNF (6 C-domain) receptor activity; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 3.
DR SMART; SM00208; TNFR; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
KW Receptor; Antigen; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 28
FT CHAIN 29 277
FT DOMAIN 29 214
FT TRANSMEM 215 235
FT DOMAIN 236 277
FT CYTOPLASMIC (POTENTIAL).
FT REPEAT 30 65
FT REPEAT 36 107
FT REPEAT 108 126
FT REPEAT 127 167
FT DISULFID 31 42
FT DISULFID 43 56
FT DISULFID 46 64
FT DISULFID 67 81
FT DISULFID 84 99
FT DISULFID 87 107
FT DISULFID 109 125
FT DISULFID 128 141
```

Matches	57;	Conservative	26;	Mismatches	88;	Indels	77;	Gaps	13														
QY	66	PWCYTTPDCVRWEYNCLTQCSETESGVLEPTPVPVPSKRAHSAAPTBOITPVPRQCYPHG	125	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :														
Dd	88	PTYLTQNPLGIGSFNL-----SEWTSQPATTPALT-TPSAES-----TPQSITTVK----	132	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :														
QY	126	NGOSYRGSTETVTGRTCOSWSNSTPHRHORTPENPNDDL--TMNYCRNPDAADTGPGWC	182	: : : :	: : :	:	:	:	:														
Dd	133	-----TKNTTTI---QLOP-SKPTTKQONKKPPKNPNHHFEVFNVPCSCSNFTC	182	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :														
QY	183	F-----TTDPSIRWEYNCLRCSDETEGTCTVVAPPTVIQVSPISGPPE	223	:	:	:	:	:	:														
Dd	183	WAICKRIPNKPCKKITPKTKPTIK-----TTKKDLKPOQTTPKAEXLTT----KETE	232	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :														
QY	224	QDCMFGNGSGKGKATT-VDTGPCOEAWAQBPBRHSSTFIPGTHKWAGLEKNVCNRPDGD	282	: : : :	: : :	:	:	:	:														
Dd	233	KPIINTKTNIIRTLLTLTNTTGNP-EITYSQKETUHLST-----SPEGN	273	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :														
QY	283	INGPWCVT 290		:	:	:	:	:	:														
Dd	274	PSFSQVYT 281		I : I																			
 RESULT 4 <hr/>																							
ID	GSHE_MOUSE	STANDARD;	PRT; 221 AA.																				
AC	P21765;																						
DT	01-MAY-1991 (Rel. 18, Created)																						
DT	01-AUG-1992 (Rel. 23, Last sequence update)																						
DT	15-SEP-2001 (Rel. 42, Last annotation update)																						
DE	Epididymal secretory glutathione peroxidase precursor (BC 1.11.1.9)																						
DE	(Epididymis-specific glutathione peroxidase-like protein). (EGLP).																						
DE	(Major androgen-regulated protein) (ARMEP24).																						
GN	gpX5.																						
QS	Mus musculus (Mouse).																						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																						
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.																						
OX	NCBII_TaxID=10090;																						
[1]																							
RZ	SEQUENCE FROM N.A.																						
RX	TISSUE=Epididymis;																						
RX	MEDLINE=G92004864; PubMed=1913244;																						
EA	Chyselneck N.B., Rigaudiere N., Dufaure J.-P.;																						
RT	'Androgen-dependent protein secreted by mouse caput epididymis shows																						
RT	high homologies with different glutathione peroxidases.";																						
RN	C. R. Acad. Sci., III, Sci. Vie 313:1-6(1991).																						
[2]																							
SOURCE OF 47-221 FROM N.A.																							
TISSUE=Epididymis;																							
MEDLINE=G1088305; PubMed=2263479;																							
CHYSSELINCK N.B., DUFAURE J.-P.																							
"A mouse cDNA sequence for epididymal androgen-regulated proteins related to glutathione peroxidase.";																							
Nucleic Acids Res. 18:7144-7144(1990).																							
-! FUNCTION:																							
Protects cells and enzymes from oxidative damage, by catalyzing the reduction of hydrogen peroxide, lipid peroxides and organic hydroperoxide, by glutathione. May constitute a glutathionine peroxidase-like protective system against peroxide damage in sperm membrane lipids.																							

```

DR EMBL; M68896; AAA37729.1; -
DR EMBL; X53780; CAA37796.1; -
DR PIR; A47367; A47367.
DR HSP; P00435; IGP1.
DR MGI; MGI:104886; Gpx5.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF00255; GSHPx; 1.
DR PRINTS; PR01011; GLUTPROXDAE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.
DR PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
KW Oxidoreductase; Peroxidase; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 221 EPIDIDYMAL SECRETORY GLUTATHIONE
      PEROXIDASE.
FT ACT_SITE 73 73 BY SIMILARITY.
FT CONFLICT 53 53 D -> H (IN REF. 2).
SQ SEQUENCE 221 AA; DESF8BD6CD22D6F9 CRC64;

Query Match 4.7%; Score 84.5; DB 1; Length 221;
Best Local Similarity 25.5%; Pred No. 5.6;
Matches 28; Conservative 10; Mismatches 39; Indels 33; Gaps 4;

QY 193 YCNLRCSDEGTGVAPPVTVIOVPSLGPSPQDCMFGNGKYGKKAFTVGTGTCQSWAA 252
DQ 72 YCGLT-----IQVELNALQEDLKPFQ-----IIVLGFPCNQGFGK 106
QY 253 QEPHRSHTFIP---TNKWAGLEKNYCRNPQDINGFWCYTMNPKLEDY 299
DQ 107 QEPGDNLEILGLKYVRGKGLFNLQFLXAGDVNGE-----NEQKIFTF 151

RESULT 5
VGLG_HRSV6 STANDARD; PRT; 297 AA.
AC P27025;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb2656).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11256;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC PIR; J01208; J01208.
DR InterPro; IPR000925; Glycoprot_G.
DR Pfam; PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA; 32708 MW; 6781756C38B6A80 CRC64;

Query Match 4.7%; Score 83.5; DB 1; Length 297;

```

```

Best Local Similarity 22.6%; Pred. No. 9.3;
Matches 56; Conservative 26; Mismatches 89; Indels 77; Gaps 13;

QY 66 PWCITTDPCVRWEYCNLCQSETSGVLETPVVPVPSMEAHSEAPTEQFPVVRQCIHG 125
DQ 88 PLYLTQNPGLGISFNSNL---SETTS---OFTTTPAPTTPS---AESTPQSTTVK----- 132
QY 126 NGQSYRGFTSTVTGRTQCSWSSMTPHRRHQTPEHPNDGL---TNMYCRNPADATGPMC 182
DQ 133 -----TKNITTT---QIOP-SKPTTKQRKNKPNKPNDFEFVNFVPCSGICNNPTC 182
QY 183 F-----TTDPSIRWEYCNLCRSDTEGTGVAPPVTVIOVPSLGPSPSE 223
DQ 183 WAICKRIPNKKPGKKTTPKTKPTIK-----TTKKDLKPQTKPKVLT-----KPT 232
QY 224 QDCMFGNGKYGKKAFT---VGTGTCQSWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPQD 282
DQ 233 KPTINTTRTNRITLTLLTTTNGNP---EYTSOKETLHST-----SPEGN 273
QY 283 INGPWCYT 290
DQ 274 PSPQVTV 281

RESULT 6
GSHE_RAT STANDARD; PRT; 221 AA.
AC P30710;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Epididymal secretory glutathione peroxidase precursor (EC 1.1.1.1.9)
DE (Epididymis-specific glutathione peroxidase-like protein) (EGLP).
GN GPX5
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wistar; TISSUE=Epididymis;
RX MEDLINE=92359957; PubMed=1386734;
RA Perry A.C.F., Jones R., Nsang L.S.P., Jackson R.M., Hall L.;
RT "Genetic evidence for an androgen-regulated epididymal secretory
RT glutathione peroxidase whose transcript does not contain a
RT selenocysteine codon.";
RL Biochem J. 285:863-870(1992).
CC -1- FUNCTION: Protects cells and enzymes from oxidative damage, by
CC catalyzing the reduction of hydrogen peroxide, lipid peroxides and
CC organic hydroperoxide, by glutathione. May constitute a
CC glutathione peroxidase-like protective system against peroxide
CC damage in sperm membrane lipids.
CC -1- CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) = oxidized
CC glutathione + 2 H(2)O.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EPIDIDYMIS.
CC -1- SIMILARITY: Belongs to the glutathione peroxidase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X62404; CAA44274.1; -.
DR PIR; S24528; S24528.
DR HSP; P00435; IGP1.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF00255; GSHPx; 1.
DR PRINTS; PR01011; GLUTPROXDAE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.

```

FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	221	EPIDIDYMAL SECRETORY GLUTATHIONE
FT	PEROXIDASE			
FT	ACT_SITE	73	73	BY SIMILARITY.
FT	SEQUENCE	221 AA;	25212 MW;	32AE912ED6G7D93 CRC64;
Query Match			4.3%;	Score 77.5; DB 1; Length 221;
Best Local Similarity			23.6%;	Pred. No. 20;
Matches 26;	Conservative 11;	Mismatches 40;	Indels 33;	Gaps 4
QY	193 YCNLRCSDTSGTGVAPPVIVQVSLGPPSEODCMFGNGKYGKATVVTGTCQSWAA	252		
Db	72 YCLT-----AQVELNALQELKPYG-----LVVLGFCNFGK	106		
QY	253 QEPHRRHFTFG---TNKWLAKENYCRNPDGDINGPCWCTYMPRKLFDY	299		
Db	107 QEPGNKBILGLKYVRPGGVPFNQLFEKGVNGE-----KEQVFSF	151		
RESULT 8				
VD	VGLG_HRSV8	STANDARD;	PRT;	292 AA.
AC	P23041;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	01-AUG-1992 (Rel. 23, Last annotation update)			
DE	Major surface glycoprotein G (Attachment glycoprotein G).			
GN	G.			
OS	Human respiratory syncytial virus (subgroup B / strain 8/60).			
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;			
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.			
OX	NCBI_TaxID=11258;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=90357765; Pubmed=1697126;			
RA	Sullender W.M., Anderson K., Wertz G.W.;			
RT	"The respiratory syncytial virus subgroup B attachment glycoprotein:			
RT	analysis of sequence, expression from a recombinant vector, and			
RT	evaluation as an immunogen against homologous and heterologous			
RT	subgroup virus challenge.";			
EL	Virology 178:195-203(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=91374595; Pubmed=1895391;			
RA	Sullender W.M., Mufson M.M., Anderson L.J., Wertz G.W.;			
RT	"Genetic diversity of the attachment protein of subgroup B			
RT	respiratory syncytial viruses.";			
RT	J. Virol. 65:5425-5434(1991).			
CC	!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE			
CC	RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND			
CC	HEMAGGLUTININATING ACTIVITIES.			
CC	!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED			
CC	CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.			
CC	!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS			
CC	DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M55633; AAA47413.1; "			
DR	EMBL; M73545; AAA47408.1; "			
DR	PIR; A37077; MGNZ60.			
DR	InterPro; IPR000325; Glycoprot.G.			
DR	Pfam; PF00802; Glycoprotein_G; 1.			
DR	Transmembrane; Glycoprotein.			
KW	DOMAIN 1 37 CYTOPLASMIC (POTENTIAL) .			
FT	TRANSMEM 38 66 POTENTIAL.			
FT	DOMAIN 67 292 EXTRACELLULAR (POTENTIAL) .			

```
FT CARBOHYD 81 81 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 252 AA; 32143 MW; 8EC6C85EF057BB5 CRC64;

Query Match 4.2%; Score 75.5; DB 1; Length 292;
Best Local Similarity 19.5%; Pred. No. 39;
Matches 48; Conservative 37; Mismatches 92; Indels 69; Gaps 10;

QY 19 ISSTV-----TGRTQSSMIPHWORTPENYPNAGLTENYCRNPDSGQPCWY 70
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DQ 69 VLTAVTVQTKNTHGKNTSTVLTQVP-----PERV-----NSSKQP---T 106
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 71 TDCPCVWEYCNLTQCSSETS-----GVLETPVVPVPSMEAHSEAPTEQPVVRQC 122
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DQ 107 TTSPIHTNSATISPNKSTHHTTAQTGRITTTQNTKPSKSKNPPKPP---KDD 162
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 123 YH-----GNGSYRGTFSTVTGR-----TCQSSMTPHRTQTPENY 161
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DQ 163 YHFEVFNVPSCIGNQLCKSICKTIPSNKPKKPTKTKTKTKRDKPTAKM 222
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 162 YHGLTMYNCNP-----DADTGPCWCFWDSIRWYCNLTQCSDEGTGVVAPTVOVP 216
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DQ 223 PKKEITNPTKKPLTKTERTDSTQSTVDTITPKY-TIQQSLHSTSENTFSQTQIP 281
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 217 SLGPPS 222
   :|||
DQ 282 TASEPS 287
   :|||

RESULT 9
WAP_RABIT
ID WAP_RABIT STANDARD; PRT; 127 AA.
AC P09412;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Whey acidic protein precursor (WAP).
GN WAP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=88335562; PubMed=3411910;
RA Devinoy E., Hubert C., Schaefer E., Houdebine L.M., Kraehenbuhl J.P.;
RT "Sequence of the rabbit whey acidic protein cDNA.";
RL Nucleic Acids Res. 16:8180-8180(1988).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90301490; PubMed=2362817;
RA Thepot D., Devinoy E., Fontaine M.L., Hubert C., Houdebine L.M.;
RT "Complete sequence of the rabbit whey acidic protein gene.";
RL Nucleic Acids Res. 18:3641-3641(1990).
CC -!- FUNCTION: COULD BE A PROTEASE INHIBITOR.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: MILK-SPECIFIC; MAJOR PROTEIN COMPONENT OF MILK WHEY.
CC -!- PTM: EIGHT DISULFIDE BONDS ARE PRESENT.
CC -!- SIMILARITY: Contains 2 WAP-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52564; CAA36798.1; -
CC EMBL; X07943; CAA30764.1; -
```

```
DR PIR; S01286; S01286.
DR HSSP; O46655; 1CJH.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 2.
DR ProDom; PD001224; WAP_C; 1.
DR SMART; SM00217; WAP; 2.
DR PROSITE; PS00317; 4-DISULFIDE_CORE; 2.
KW Milk; Whey; Protease inhibitor; Repeat; Phosphorylation; Signal.
FT SIGNAL 1 19
FT CHAIN 20 127 WHEY ACIDIC PROTEIN.
FT DOMAIN 28 68 WAP 1 (ATYPICAL).
FT DOMAIN 74 121 WAP 2 (ATYPICAL).
FT MOD_RES 237 237 PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 127 AA; 13526 MW; 59DAD5FBFD2E2087 CRC64;

Query Match 4.2%; Score 75; DB 1; Length 127;
Best Local Similarity 26.2%; Pred. No. 17;
Matches 17; Conservative 9; Mismatches 31; Indels 8; Gaps 2;

QY 57 CRNPDSGQPCWYTDPCVWEYCNLTQCSSETSGLVETPTVVPVPSMEAHSEAPTEQ 116
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DQ 32 CPEFSSEETCLSDNDCLGSTVC----CPSAAGGSCRTPIIVTPPK----AGRCPWQA 83
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 117 PVVRQ 121
   :|||
DQ 84 PMLSQ 88
   :|||

RESULT 10
GSHY_RAT
ID GSHY_RAT STANDARD; PRT; 221 AA.
AC Q64625;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Glutathione peroxidase RX2D1 precursor (EC 1.11.1.9) (Odorant-
DE metabolizing protein RX2D1).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=92031476; PubMed=1931961;
RA Dear T.N., Campbell K., Rabbitts T.H.;
RT "Molecular cloning of putative odorant-binding and odorant-
RT metabolizing proteins.";
RL Biochemistry 30:10376-10382(1991).
CC -!- CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) = oxidized
CC glutathione + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE BOWMAN'S GLANDS.
CC -!- SIMILARITY: Belongs to the glutathione peroxidase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M76733; AAA42094.1; -
CC PIR; B40464; B40464.
DR HSSP; P00435; 1GPI.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF00255; GSHPx; 1.
DR PRINTS; PR01011; GLUTPOXDASE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.
DR PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
KW Oxidoreductase; Peroxidase; Signal.
FT SIGNAL 1 21
   POTENTIAL.
```


FT	CHAIN	22	221		GLUTATHIONE PEROXIDASE RY2D1.
FT	ACT_SITE	73	73		BY SIMILARITY.
SQ	SEQUENCE	221 AA;	24961 NW;	92749EAF6A3EE48C CRC64;	
	Query Match	4.2%;	Score 75;	DB 1;	Length 221;
	Best local similarity	25.0%;	Pred. NO. 31;		
Matches	23;	Conservative 13;	Mismatches 32;	Indels 24;	Gaps 3;
QY	227 MFGNGXGYRGKKAT-----	TVTGTCQEWAAQEPHRSHTFPG---	TNK 267		
	: : : : :	: : : : :	:		
Db	65 LFVNASFCGLRATYPPELNTQLBELRFNVSLGVPCNFKGQKPGKNSIILLGLKYVRP 124				
	: : : : :	: : : : :	:		
QY	268 WAGLEKNYCNPDCDINGPCWCTHNPKLFDY 199				
	: : : : :	: : : : :	:		
Db	125 GGGVPFPNFOLFERKGVNDG-----NEQKVFSF 151				
	: : : : :	: : : : :	:		
RESULT 11					
GSHE_PIG					
ID	GSHE_PIG	STANDARD;	PRG:	219 AA.	
AC	O18994;				
DT	15-JUL-1999 (Rel. 38, Created)				
DT	15-JUL-1999 (Rel. 38, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	Epididymal secretory glutathione peroxidase precursor (EC 1.11.1.9)				
DE	(Epididymis-specific glutathione peroxidase-like protein) (EGLP).				
GN	GPX5.				
OS	Sus scrofa (pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OX	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
NCBI_TaxID	=9823;				
[1]					
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 22-46.				
RP					
RC	TISSUE=Epididymis;				
RX	MEDLINE=97415436; PubMed=9271255;				
RA	Okamura N., Iwaki Y., Hiramoto S., Tamba M., Bannai S., Sugita Y.,				
RA	Syntin P., Dacheux F., Dacheux J.L.;				
RT	"Molecular cloning and characterization of the epididymis-specific				
RT	glutathione peroxidase-like protein secreted in the porcine				
RT	epididymal fluid".				
RL	Biochim. Biophys. Acta 1336:99-109(1997).				
-!	FUNCTION: MAY CONSTITUTE A GLUTATHIONINE PEROXIDASE-LIKE				
CC	PROTECTIVE SYSTEM AGAINST PEROXIDE DAMAGE IN SPERM MEMBRANE				
CC	LIPIDS. SINCE THE PURIFIED PORCINE ENZYME HAS VERY LITTLE ACTIVITY				
CC	TOWARDS HYDROGEN PEROXIDE OR ORGANIC HYDROPEROXIDES THE PROTECTIVE				
CC	EFFECT IS NOT LIKELY TO BE EXERTED BY ITS ENZYMATIC ACTIVITY.				
CC	INSTEAD, MAY PROTECT SPERM FROM PREMATURE ACROSOME REACTION IN THE				
CC	EPIDIDYMS BY BINDING TO LIPID PEROXIDES, WHICH MIGHT OTHERWISE				
CC	INTERACT WITH PHOSPHOLIPASE A2 AND INDUCE THE ACROSOME REACTION.				
-!	CATALYTIC ACTIVITY: 2 glutathione + H(2O)(2) - oxidized				
CC	glutathione + 2 H(2O).				
-!	SUBUNIT: Homotrimer.				
-!	SUBCELLULAR LOCATION: Secreted.				
CC	TISSUE SPECIFICITY: PROXIMAL CAPUT EPIDIDYMS.				
-!	SIMILARITY: Belongs to the glutathione peroxidase family.				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation-				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to licenset@isb-sib.ch).				
CC					
DR	EMBL; D37916; BAA22149.1; -				
DR	HSP; P00435; 1GP1.				
DR	InterPro; IPRO0089; Glut_peroxidase.				
DR	Pfam; PF00255; GSHPx; 1.				
DR	PRINTS; PR01011; GLUTPEROXDASE.				
DR	PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.				
DR	PROSITE; FS00763; GLUTATHIONE_PEROXID_2; 1.				
KW	Oxidoreductase; Peroxidase; Signal.				
FT	FT	1	21		

RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA	Takenuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA	Viari A., Wambut R., Wedler E., Wedler H., Weitzenegger T.,
RA	Minters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA	Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT	"The complete genome sequence of the Gram-positive bacterium <i>Bacillus</i>
RT	<i>subtilis</i> .";
RT	Nature 390:249-256(1997).
RN	[3]
RN	SEQUENCE OF 1-71 FROM N.A.
RC	STRAIN=168 / BGSC1A1;
RC	MEDLINE=96093926; PubMed=7584049;
RY	Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT	"Cloning and sequencing of a 36-kb region of the <i>Bacillus subtilis</i>
RT	genome between the gnt and iol operons.";
RT	DNA Res. 2:61-69(1995).
CC	!- PATHWAY: Myo-inositol catabolism.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; D14399; BAA03291.1; -;
DR	EMBL; 299124; CAB16011.1; -;
DR	EMBL; AB005554; BAA21610.1; -;
DR	PIR; B69645; B69645.
DR	Subtilist; BG11118; iolB.
KW	Complete proteome.
QY	SEQUENCE 271 AA; 30770 MW; 4322AAE4C18DF21F3 CRC64;
QY	Query Match 4.1%; Score 74; DB 1; Length 271;
Db	Best Local Similarity 20.3%; Pred. No. 47;
Db	Matches 42; Conservative 19; Mismatches 64; Indels 82; Gaps 10;
QY	101 VFSMEAHSEAAPEQFPV-----VRCY-----HGNGOSYR 131
Db	76 IGTRESYPERKPTDSVLIISNDFEITAVSDARVALCYSPSEKQLPTKLIKAEDNGIEHR 135
QY	132 GFSTTVTGR-----CQSWSSMTFHRHQRTPENYPNDG-LTMAY 170
Db	136 GQFSNKRPHVNLPSDSDPSANSLWVEVTDGNNWSPPHKEDQ--NLNPEESFLEET 193
QY	171 CRNPADAGT---PWCFTTDPISIRWEYNLTRCDTGEVWVAPTVQPS-----LGPPE 223
Db	194 YHELDPGGFFQFVYTDRSI-----DEMTVGNENVVIPAGYHPVGVPD- 240
QY	224 QDCMFGNGKGVKSKATVTGTPCQEW 250
Db	241 -----GYTSYILNNMAG-PTRKW 257
RESULT 14	
VGIG_HRSV1	
ID	VGIG_HRSV1 STANDARD; PRT; 292 AA.
AC	P20896;
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, Last sequence update)
DT	01-NOV-1991 (Rel. 20, Last annotation update)
DE	Major surface glycoprotein G (Attachment glycoprotein G).
OS	G.
GN	Human respiratory syncytial virus (subgroup B / strain 18537).
OC	Viruses; sRNA negative-strand viruses; Mononegavirales;
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX	NCBI_TaxID=11251;
RN	[1]
RN	SEQUENCE FROM N.A.
RX	MEDLINE=8728957; PubMed=2441388;
RX	Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;

RT "The G glycoprotein of human respiratory syncytial viruses of
 RT subgroups A and B: extensive sequence divergence between
 RT antigenically related proteins"; J. Virol. 56:2525-2529 (1987).
 RL Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629 (1987).
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
 CC HEMAGGLUTININATING ACTIVITIES.
 CC
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
 CC
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; M17213; AAA47412.1; -
 DR PIR; B32703; MGN218.
 DR InterPro; IPR000925; Glycoprot G.
 DR Pfam; PF00802; Glycoprotein_G; 1.
 KW Transmembrane; Glycoprotein.
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 38 66 POTENTIAL.
 FT DOMAIN 67 292 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 292 AA; 32306 MW; BC8C59F69CA7AFC2 CRC64;
 Query Match 4.1%; Score 73.5; DB 1; Length 292;
 Best Local Similarity 20.6%; Pred. No. 56;
 Matches 49; Conservative 36; Mismatches 98; Indels 53; Gaps 9;
 QY 19 ISSTVTGRTQSSSMIPHHQHTPTENPNAGLLENYCRNPDGKOPWCYTDFCVKWE 78
 DB 69 VITLITVQTIKN-----HTKNI-STYLQVPPERNSSKQP---TITSPITN 114
 QY 79 YCNLTQCSFES-----GVLETFVVPVPSMEAHSEAPTEQTQPVVRCQYH----- 124
 DB 115 SATISPTNCEGHTHTAQTGRITTTQTNKPSKSNPKKP---KDDYHFEVNF 170
 QY 125 -----GNCQSRGTFTVTGR-----TCQSSSMTPHHQHTPTENPNAGLTFN 169
 DB 171 VPCISGNNOLCKSICKTIPSNKPKKPTIKPTNKPTTKTNKDPKPAKPKKEIITN 230
 QY 170 YCRNP-----DADTGFWCFFTDFPSIRWEYCNLTGCSDTGTVVAPTVIQVPSLGPSS 222
 DB 231 PAKKPKTKTTRDTSISQSTVLDTITPKY-TIQQQLSHSTTSNTSPSTQIPTASEPS 287
 RESULT 15
 KNHL_CANGA
 ID KNHL_CANGA STANDARD; PRT; 265 AA.
 AC O74684;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Cell wall synthesis protein KNHL precursor.
 GN KNHL.
 OS Candida glabrata (Yeast) (Torulopsis glabrata).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5478;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 2001;
 RX MEDLINE=98422452; PubMed=9748432;
 RA Nagahashi S., Iussier M., Bussey H.;
 RT "Isolation of Candida glabrata homologs of the Saccharomyces

RT cerevisiae KRE9 and KNH1 genes and their involvement in cell wall
 RT beta-1,6-glucan synthesis."; J. Bacteriol. 180:5020-5029 (1998).
 RL J. Bacteriol. 180:5020-5029 (1998).
 CC -!- FUNCTION: INVOLVED IN CELL WALL BETA(1->6) GLUCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: SECRETED, PROBABLY FOUND AT CELL SURFACE (BY
 CC SIMILARITY).
 CC -!- PTM: O-GLYCOSYLATED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE KRE9 / KNH1 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; AF064252; AAC64009.1; -
 KW Glycoprotein; Cell wall; Signal.
 FT SIGNAL 1 14 POTENTIAL.
 FT CHAIN 15 265 CELL WALL SYNTHESIS PROTEIN KNH1.
 SQ SEQUENCE 265 AA; 29046 MW; 6A13BCD6BC0AE2BC CRC64;
 Query Match 4.1%; Score 73; DB 1; Length 265;
 Best Local Similarity 20.1%; Pred. No. 55;
 Matches 57; Conservative 38; Mismatches 88; Indels 100; Gaps 16;
 QY 52 LLENYCRNPDGKQP--WCYTDFPCVWEYCNLT--OCSETESGVLETFVVPVPSMEAH 107
 DB 23 LDQSFEDASGGTAKIPTQWLFTPTNPQDFTSLTFLSCG-----PNYKIEAF 70
 QY 108 -----SEAAPTQTPVVRQCYHGNGOSYRGFTSTVTGTCQSSSMTPHHQHTPTENY 161
 DB 71 KVIGKLSDTGTFDEAEVSQSVGANGYIVQITAA----- 106
 QY 162 PNDGLTMYCRNPDADTGPWCFTTDPDSIRWEYCNLT--RCSDEGTGVVAPTV---IQVP 216
 DB 107 -TDGTTIHY-----SPRKLTCMTGSKLPDIL-LITAPPTPTETRTTG 147
 QY 217 SLGPPSEQ---DCMFG--NGKG---YEGKATVTVGTCQEWAAQ----- 253
 DB 148 DLGATIDSKSFDPYGEONGKAKFAPMOTQPGKITAT---TWSRYATSASFSTSLTA 204
 QY 254 EPHRSTFIPGTNKGWGLEKNYC-RNPDGDLINGPWCYTNPKR 295
 DB 205 TPVQHTLTFGWSYIISDYNYAPAPPPSDNGW---YDPKK 244
 Search completed: September 29, 2003, 08:02:14
 Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:01:08 ; Search time 36 Seconds
(without alignments)
2207.783 Million cell updates/sec

Title: US-10-088-548-2

Perfect score: 1786

Sequence: 1 KSPVVQDCYHGDSRGYS.....YTMNPKLFYCDIFLCASS 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 330525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 513101

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhcc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	644	36.1	113	4	Q9UIR5	Q9uir5 homo sapien
2	634	35.5	113	4	Q9UIR6	Q9uir6 homo sapien
3	550.5	30.8	215	13	Q42341	Q42341 gallus gall
4	544	30.5	95	4	Q8N696	Q8n696 homo sapien
5	498	27.9	132	4	Q16609	Q16609 homo sapien
6	489	27.4	113	4	Q9UIR7	Q9uir7 homo sapien
7	481.5	27.0	290	4	Q02935	Q02935 homo sapien
8	477.5	26.7	296	4	Q14519	Q14519 homo sapien
9	467.5	26.2	145	6	Q28911	Q28911 macaca fasc
10	466	26.1	285	4	Q8TCE2	Q8tce2 homo sapien
11	445	24.9	105	4	Q9UIR8	Q9uir8 homo sapien
12	261.5	14.6	109	6	Q9N1B8	Q9n1b8 ovis aries
13	260	14.6	209	11	Q8B517	Q8bs17 mus musculus
14	249.5	14.0	208	4	Q8BYN0	Q8byn0 homo sapien
15	249.5	14.0	210	4	Q13494	Q13494 homo sapien
16	248	13.9	53	4	Q9UMI2	Q9um12 homo sapien

17	243.5	13.6	211	11	Q55027	O55027 mus musculus
18	241	13.5	60	4	Q9UKJ7	Q9ukj7 homo sapien
19	231.5	13.0	90	4	Q8NG20	Q8ng20 homo sapien
20	225	12.6	202	13	Q90675	Q90675 gallus gall
21	210.5	11.8	75	6	Q9EGN9	Q9egn9 bos taurus
22	168.5	9.4	111	6	Q77688	Q77688 cryptolagus
23	165.5	9.3	157	6	Q9TVA8	Q9tva8 bos taurus
24	159	8.9	231	11	Q8C6L2	Q8c6l2 mus musculus
25	152	8.5	154	4	Q96SE8	Q96se8 homo sapien
26	141	7.9	25	4	Q9UD88	Q9ud88 homo sapien
27	117	6.6	263	4	Q00318	O00318 homo sapien
28	117	6.6	263	4	Q96FE7	Q96fe7 homo sapien
29	117	6.6	263	4	Q8NCJ9	Q8ncj9 homo sapien
30	114	6.4	162	11	Q8C4E2	Q8c4e2 mus musculus
31	114	6.4	286	4	Q13209	Q13209 homo sapien
32	95.5	5.3	214	6	Q9XT70	Q9xt70 cryptolagus
33	91	5.1	261	5	Q9VAX5	Q9vax5 drosophila
34	90.5	5.1	268	10	Q42830	Q42830 phytophthor
35	89.5	5.0	289	16	Q8ZJU7	Q8zju7 salmonella
36	88.5	5.0	289	16	Q8ZOT3	Q8zot3 salmonella
37	87.5	4.9	223	4	Q96RT0	Q96rt0 homo sapien
38	87	4.9	249	10	Q24102	Q24102 medicago tr
39	86.5	4.8	223	4	Q96BH3	Q96bh3 homo sapien
40	86.5	4.8	295	12	Q86361	Q86361 respiratory
41	85.5	4.8	272	16	Q8YFU6	Q8yfu6 brucella me
42	85	4.8	234	11	Q9CUJ3	Q9cu13 mus musculus
43	84.5	4.7	220	4	Q9H4C8	Q9h4c8 homo sapien
44	84.5	4.7	221	11	Q8CDQ5	Q8cdq5 mus musculus
45	84.5	4.7	241	10	Q9SUP1	Q9sup1 arabidopsis

ALIGNMENTS

RESULT 1

Q9UIR5 PRELIMINARY; PRT; 113 AA.

AC Q9UIR5;

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Apolipoprotein(a) (Fragment).

GN APOA.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21181705; PubMed=11285247;

RA Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;

RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV types 6 to 10 domain affect Lp(a) plasma concentrations and have different patterns in Africans and Caucasians.";

RL Hum. Mol. Genet. 10:815-824(2001).

CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AF158663; AAF03680.1; -

DR HSSP; P00747; LPMK.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF000511; kringle; 1.

DR PRINTS; PR00018; KRINGLE.

DR PRODOM; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

KW Glycoprotein; Kringle; Lipoprotein.

FT NON_TER 1

FT NON_TER 113

SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;

Query Match 36.1%; Score 644; DB 4; Length 113;

Best Local Similarity 99.1%; Pred. No. 1.2e-49;

Matches 112; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 112 PTEQTPVRCYHGNGQSYRGFTFTVTGRTCSQSSMTPHRRHQTPTNPGLTMYTC 171
 DB 1 PTEQTPVRCYHGNGQSYRGFTFTVTGRTCSQSSMTPHRRHQTPTNPGLTMYTC 60

QY 172 RNPADTGPWCFTTDPISIRWEYCNLTQCSOTEGTVVAPPTVIOVPSLGPPEQ 224
 DB 61 RNPADTGPWCFTTDPISIRWEYCNLTQCSOTEGTVVAPPTVIOVPSLGPPEQ 113

RESULT 2

Q9UR6 PRELIMINARY; PRT; 113 AA.

AC Q9UR6; PRT; 113 AA.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Apolipoprotein(a) (Fragment).

GN APOA.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RA [1]

RP SEQUENCE FROM N.A.

RE MEDLINE=21181705; PubMed=11285247;

RA Ohtsuka M., Kraft H.G., Ehmholm C., Utermann G.;
 RT "Single nucleotide polymorphisms in exons of the apo(a) kringle IV
 RT types 5 to 10 domain affect Lp(a) plasma concentrations and have
 RT different patterns in Africans and Caucasians.";
 RL Hum. Mol. Genet. 10:815-824(2001).

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AF158661; AAF03679.1; -;
 DR EMBL; AF158660; AAF03679.1; JOINED.

DR HSP; P00747; 2PK4.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; kringle; 1.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

KW Glycoprotein; Kringle; Lipoprotein.

FT NON_TER 1 113
 FT TER 113 113
 SQ SEQUENCE 113 AA; 12697 MW; 51D4461D9C66312E CRC64;

Query Match 35.5%; Score 634; DB 4; Length 113;
 Best Local Similarity 100.0%; Pred. No. 9.2e-49;
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPVQDCYHGDSYRGISSTVTGRTCSQSSMTPHRRHQTPTNPGLTMYTCRNP 60
 DB 4 KSPVQDCYHGDSYRGISSTVTGRTCSQSSMTPHRRHQTPTNPGLTMYTCRNP 63

QY 61 DSGKQPCWYTTDPCVRWEYCNLTQCSQSSMTPHRRHQTPTNPGLTMYTCRNP 110
 DB 64 DSGKQPCWYTTDPCVRWEYCNLTQCSQSSMTPHRRHQTPTNPGLTMYTCRNP 113

RESULT 3

O42341 PRELIMINARY; PRT; 215 AA.

AC O42341;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE HGF alpha-chain (Fragment).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Heart;

RA Isokawa K., Yahagi N., Honma J., Tanaka M., Murakami K., Yokoyama H.,
 RA Toda Y.;

RT "The expression of hepatocyte growth factor mRNA is temporally
 RT relevant to cardiac endothelial-mesenchymal transformation.";
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 3 KRINGLE DOMAINS.

DR EMBL; D63779; BAA23643.1; -;
 DR HSP; P00747; IKN.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; kringle; 3.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 3.

DR SMART; SM00130; KR; 2.

DR PROSITE; PS00021; KRINGLE_1; 3.

DR PROSITE; PS00070; KRINGLE_2; 3.

KW Glycoprotein; Kringle.

FT NON_TER 1 215
 FT TER 215 215
 SQ SEQUENCE 215 AA; 24949 MW; 55E67AB52DAF316 CRC64;

Query Match 30.8%; Score 550.5; DB 13; Length 215;
 Best Local Similarity 40.9%; Pred. No. 4.9e-41;
 Matches 104; Conservative 24; Mismatches 81; Indels 45; Gaps 6;

QY 44 PENYPNAGLTENYCRNPDSGKQPCWYTTDPCVRWEYCNLTQCSQSSMTPHRRHQTPTNPGLTMYTCRNP 103
 DB 2 PERPDGFDNDYCRNPDSGKQPCWYTTDPCVRWEYCNLTQCSQSSMTPHRRHQTPTNPGLTMYTCRNP 50

QY 104 MEAHSEAAPEQTPVRCYHGNGQSYRGFTFTVTGRTCSQSSMTPHRRHQTPTNPGLTMYTCRNP 163
 DB 51 -----TEAVAETTCIOGQSGYRGFTFTVTGRTCSQSSMTPHRRHQTPTNPGLTMYTCRNP 101

QY 164 DGLTMVYCRNPADTGPWCFTTDPISIRWEYCNLTQCSQSSMTPHRRHQTPTNPGLTMYTCRNP 223
 DB 102 KDLRENYCRNPDSGKQPCWYTTDPCVRWEYCNLTQCSQSSMTPHRRHQTPTNPGLTMYTCRNP 143

QY 224 QDCMFGNGKGYRGKATTTGTGTPQEW--AAQEPHRRHSTFI--PGTNKRWAGLEKNTCRNP 279
 DB 144 QDCYRGNGKSYNGNLSTKTEGLTCTWDKNIEDLRHQTPTNPGLTMYTCRNP 200

QY 280 DGDINGPWCYTMNP 293
 DB 201 DDDSGPWCYTDTP 214

RESULT 4

O8N696 PRELIMINARY; PRT; 95 AA.

AC O8N696;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Antiangiogenic AK38 protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Dou D.;

RT "Antiangiogenic domain of apolipoprotein(a).";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AY039748; AAK74187.1; -;

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003966; Prothrombin.

DR Pfam; PF00051; kringle; 1.

DR PRINTS; PR00018; KRINGLE.

```

DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Glycoprotein; Kringle.
SQ SEQUENCE 95 AA; 10553 MW; 31EDC9DB38A372DF CRC64;

Query Match 30.5%; Score 544; DB 4; Length 95;
Best Local Similarity 98.9%; Pred. No. 7.1e-41;
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 215 VPSLGPSPDQCMFGNGKRGKATVGTGFCQWAAQEPHRSFPIGTNKGWLEKN 274
DQ 3 VPSLGPSPDQCMFGNGKRGKATVGTGFCQWAAQEPHRSFPIGTNKGWLEKN 62
QY 275 YCNRPDGLNGPWCYTMNPKLFDYCDIPLCA 307
DQ 53 YCNRPDGLNGPWCYTMNPKLFDYCDIPLCA 95

RESULT 5
Q16609
ID Q16609 PRELIMINARY; PRT; 132 AA.
AC Q16609;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE (APOARGC).
GN APOARGC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=95268939; PubMed=7749817;
RA Byrne C.D., Schwartz K., Lawn R.M.;
RT "Loss of a splice donor site at a 'skipped exon' in a gene homologous
RT to apolipoprotein(a) leads to an mRNA encoding a protein consisting of
RT a single kringle domain."
RL Arterioscler. Thromb. Vasc. Biol. 15:65-70(1995).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; U19518; AAA85693.1; -.
DR EMBL; U19517; AAA85692.1; -.
DR HSP; P00747; IPMK.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Glycoprotein; Kringle.
SQ SEQUENCE 132 AA; 14886 MW; 3794AD30A586DBBA CRC64;

Query Match 27.9%; Score 498; DB 4; Length 132;
Best Local Similarity 75.5%; Pred. No. 1.2e-36;
Matches 83; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 109 EAATGTFPVRCCHGNGSYRTFTTGTGTCQSWSSMTPHRHORTPENPNDGLTM 168
DQ 15 KSAPTETGSPVQECYHSNGSYRTFTTGTGTCQSWSSMTPHRHORTPENPNDGLIS 74
QY 169 YCNRPDADTGPWCFTTDPNVRWEYCNLTRCSDTGTCTVAPPTVQVPSL 218
DQ 75 YCNRPDADTGPWCFTTDPNVRWEYCNLTRCSDTGTCTVAPPTVQVPSL 124

RESULT 6
Q9UIR7
ID Q9UIR7 PRELIMINARY; PRT; 113 AA.

```

```

AC Q9UIR7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=21181705; PubMed=11285247;
RA Ogorekova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect lip(a) plasma concentrations and have
RT different patterns in Africans and Caucasians."
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158659; AAF03678.1; -.
DR EMBL; AF158658; AAF03678.1; JOINED.
DR HSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Glycoprotein; Kringle; Lipoprotein.
KW NON_TER 1
FT NON_TER 113
FT SEQUENCE 113 AA; 12815 MW; 4F80ADF8708548CB CRC64;

Query Match 27.4%; Score 489; DB 4; Length 113;
Best Local Similarity 78.0%; Pred. No. 6.5e-36;
Matches 85; Conservative 5; Mismatches 19; Indels 0; Gaps 0;

QY 1 KSPVVQDCYHGDGSRYSRISSTVTGRTCSQSSMIPHHORTPENPAGLTENCRNP 60
DQ 4 QSPVVQDCYHGDGSRYSRISSTVTGRTCSQSSMIPHHORTPENPAGLTENCRNP 63
QY 61 DSGKQPCWCTTDPNVRWEYCNLTRCSDTGTCTVAPPTVQVPSMEHSE 109
DQ 64 DREIRPWCYTMDSVRWEYCNLTRCSDTGTCTVAPPTVQVPSMEHSE 112

RESULT 7
Q02935
ID Q02935 PRELIMINARY; PRT; 290 AA.
AC Q02935;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hepatocyte growth factor, heavy chain precursor.
GN HGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=91200041; PubMed=1826653;
RA Miyazawa K., Kitamura A., Naka D., Kitamura N.;
RT "An alternatively processed mRNA generated from human hepatocyte
RT growth factor gene."
RL Eur. J. Biochem. 197:15-22(1991).
RN [2]
SEQUENCE FROM N.A.
MEDLINE=93087571; PubMed=1280830;
RA Hartmann G., Naldini L., Weidner K.M., Sachs M., Vigna E.,
RA Comoglio P.M., Birdmeier W.;

```

RT "A functional domain in the heavy chain of scatter factor/hepatocyte growth factor binds the c-Met receptor and induces cell dissociation but not mitogenesis.";

RT Proc. Natl. Acad. Sci. U.S.A. 89:11574-11578(1992).

CC -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES. IT HAS NO DETECTABLE PROTEASE ACTIVITY.

CC -!- SIMILARITY: CONTAINS TWO KRINGLE REGIONS.

CC -!- SIMILARITY: HIGH, TO OTHER HGF; LOWER, TO PLASMINOGEN.

DR EMBL; X57574; CAA40802.1; -

DR EMBL; L02931; AAA52649.1; -

DR HSSP; P14210; 1BHT.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan_app.

DR Pfam; PF00051; Kringle; 2.

DR Pfam; PF00024; PAN; 1.

DR PRINTS; PR00018; KRINGLE.

DR PRODOM; PD000395; Kringle; 2.

DR SMART; SM00473; PAN.AP; 1.

DR PROSITE; PS00021; KRINGLE_1; 2.

DR PROSITE; PS00070; KRINGLE_2; 2.

DR GlycoProtein; Kringle.

SW SEQUENCE 296 AA; 33765 MW; C8A18A6FD063200A CRC64;

Query Match 27.08; Score 481.5; DB 4; Length 290;

Best Local Similarity 43.74; Pred. No. 9e-35;

Matches 86; Conservative 21; Mismatches 55; Indels 35; Gaps 3;

QY 5 VDCYHGDSYRGISSTVTGTCQSSSMIPHWQRTPTNPAGLTENCRNPDGSK 64

DB 125 INRCIIIGKRSYKGTIVSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNP-RGE 183

QY 65 Q--PCYTTDPCVRYEYCNLTQCSFESGVLETPVTVVPSMEAHSEAAPTQTPVVRQC 122

DB 184 EGGPWCFSTSNPEVRYEYCDIPQCSVE-----C 211

QY 123 YHNGQSYRGFTSTVTGTCQSSSMTPRHQRTPTNPNDGLTWNVYCRNPDADTGPWC 182

DB 212 MTCNGESYRGMDHTSGIKQCRWDHQTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 271

QY 183 FTTDPSIRWEYCNLTQCS 199

DB 272 YTLDPHTRWEYCAIKTC 288

RESULT 8

Q14519 PRELIMINARY; PRT; 296 AA.

AC Q14519;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Competitive HGF antagonist.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RA Chan A.M.-L.; Rubin J.S.; Bottaro D.P.; Hirschfield D.W.; Chedid M.; Aaronson S.A.;

RA "Identification of a Competitive HGF Antagonist Encoded by an Alternative Transcript.";

RT

RL Science 0:0-0(1991).

CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

DR EMBL; M7727; AAA35980.1; -

DR HSSP; P14210; 1BHT.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan_app.

DR Pfam; PF00051; Kringle; 2.

DR Pfam; PF00024; PAN; 1.

DR PRINTS; PR00018; KRINGLE.

DR PRODOM; PD000395; Kringle; 2.

DR SMART; SM00130; KR; 2.

DR SMART; SM00473; PAN.AP; 1.

DR PROSITE; PS00021; KRINGLE_1; 2.

DR PROSITE; PS00070; KRINGLE_2; 2.

DR GlycoProtein; Kringle.

SW SEQUENCE 296 AA; 34546 MW; A45E456B87AE03BE CRC64;

Query Match 26.7%; Score 477.5; DB 4; Length 296;

Best Local Similarity 43.28; Pred. No. 2.1e-34;

Matches 86; Conservative 21; Mismatches 57; Indels 35; Gaps 3;

QY 5 VDCYHGDSYRGISSTVTGTCQSSSMIPHWQRTPTNPAGLTENYCRNPDGSK 64

DB 125 INRCIIIGKRSYKGTIVSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNP-RGE 183

QY 65 Q--PCYTTDPCVRYEYCNLTQCSFESGVLETPVTVVPSMEAHSEAAPTQTPVVRQC 122

DB 184 EGGPWCFSTSNPEVRYEYCDIPQCSVE-----C 211

QY 123 YHNGQSYRGFTSTVTGTCQSSSMTPRHQRTPTNPNDGLTWNVYCRNPDADTGPWC 182

DB 212 MTCNGESYRGMDHTSGIKQCRWDHQTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 271

QY 183 FTTDPSIRWEYCNLTQCS 201

DB 272 YTLDPHTRWEYCAIKNRD 290

RESULT 9

Q28911 PRELIMINARY; PRT; 145 AA.

AC Q28911;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Apolipoprotein[a] (Fragment).

GN APOLIPOPROTEIN[A].

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopithecinae; Macaca.

OC NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=95395387; PubMed=7666007;

RA Ramharack R.; Spahr M.A.; Hicks G.W.; Kieft K.A.; Brammer D.W.; Minton L.L.; Newton R.S.;

RA "Gemfibrozil significantly lowers cynomolgus monkey plasma lipoprotein[a]-protein and liver apolipoprotein[a] mRNA levels.";

RL J. Lipid Res. 36:1294-1304(1995).

CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; S79621; AAD14312.1; -

DR HSSP; P00747; 2PK4.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; Kringle; 1.

DR PRINTS; PR00018; KRINGLE.

DR PRODOM; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR GlycoProtein; Kringle; Lipoprotein.

SW NON_TER 145 145

```
SQ SEQUENCE 145 AA; 16097 MW; 0D2DA20A06A5DD87 CRC64;
Query Match 26.2%; Score 467.5; DB 6; Length 145;
Best Local Similarity 64.7%; Pred. No. 7e-34;
Matches 86; Conservative 11; Mismatches 25; Indels 11; Gaps 2;

QY 102 PSMEHSE-----AAPTEQTPVVRQCYHGNGSYRGFTFTVTGRTQSSWMT 151
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 14 PIME-HKEVPLLLFLKALPEQSHVVDQCYHGDQSYGCTSTTGTGTCQAWSSMEP 72

QY 152 HRHQRPEVNDGLTMKNCNPDADTGPWCETTPDSIRWEYCNLTGCDTGTGTVAPPT 211
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 73 HOHNRTEYNPAGIRNCRNPDVPAAPYCYTMDPNVRWEYCNLTGCDTGTGTVAPPT 132

QY 212 VIOVPSLGPPSPSQ 224
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 133 VTLVPSLEAPSPSQ 145

RESULT 10
Q8TCE2 PRELIMINARY; PRT; 285 AA.
AC Q8TCE2
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to hepatocyte growth factor (hepatopoietin A, scatter factor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorolkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringle IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AFI58656; AAF03677.1; -.
DR HSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; Kringle.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11882 MW; 6ECB6C02CD30EFA2 CRC64;
Query Match 24.9%; Score 445; DB 4; Length 105;
Best Local Similarity 72.5%; Pred. No. 4.7e-32;
Matches 74; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 112 PTEQPVVRQCYHGNGSYRGFTFTVTGRTQSSWMTPHRHQRTPEYNDGLTMNYC 171
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1 PTEQSPGVQDCYHGDQSYRGFTFTVTGRTQSSWMTPHRHQRTPEYNDGLTMNYC 60

QY 172 RNPDAATGTCWCTTDPDSIRWEYCNLTGCDTGTGTVAPPTVI 213
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 61 RNPDAEISFWCTMDPNVRWEYCNLTGCDTGTGTVAPPTVI 102

RESULT 12
Q9N1B8 PRELIMINARY; PRT; 109 AA.
AC Q9N1B8
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hepatocyte growth factor (Fragment).
GN HGF.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Endometrium;
RX MEDLINE=20297031; PubMed=10819791;
RA Chen C., Spencer T.E., Bazer F.W.;
RT "Expression of hepatocyte growth factor and its receptor c-met in the
RT ovine uterus.";
RL Biol. Reprod. 62:1844-1850(2000).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
```

```
RESULT 11
Q9UIR8 PRELIMINARY; PRT; 105 AA.
AC Q9UIR8
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorolkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AFI58656; AAF03677.1; -.
DR HSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; Kringle.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11882 MW; 6ECB6C02CD30EFA2 CRC64;
Query Match 24.9%; Score 445; DB 4; Length 105;
Best Local Similarity 72.5%; Pred. No. 4.7e-32;
Matches 74; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 112 PTEQPVVRQCYHGNGSYRGFTFTVTGRTQSSWMTPHRHQRTPEYNDGLTMNYC 171
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1 PTEQSPGVQDCYHGDQSYRGFTFTVTGRTQSSWMTPHRHQRTPEYNDGLTMNYC 60

QY 172 RNPDAATGTCWCTTDPDSIRWEYCNLTGCDTGTGTVAPPTVI 213
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 61 RNPDAEISFWCTMDPNVRWEYCNLTGCDTGTGTVAPPTVI 102

RESULT 12
Q9N1B8 PRELIMINARY; PRT; 109 AA.
AC Q9N1B8
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hepatocyte growth factor (Fragment).
GN HGF.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Endometrium;
RX MEDLINE=20297031; PubMed=10819791;
RA Chen C., Spencer T.E., Bazer F.W.;
RT "Expression of hepatocyte growth factor and its receptor c-met in the
RT ovine uterus.";
RL Biol. Reprod. 62:1844-1850(2000).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
```



```

DR EMBL; AF213397; AAF25945.1; -.
DR HSP; P14210; 2HGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF000051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Glycoprotein; Kringle.
KW NON_TER 1
FT NON_TER 1
SQ SEQUENCE 109 AA; 1F88FE5DBC0D4A5D CRC64;

Query Match 14.6%; Score 261.5; DB 6; Length 109;
Best Local Similarity 39.0%; Pred. No. 9.1e-16;
Matches 48; Conservative 15; Mismatches 37; Indels 23; Gaps 2;

Qy 114 EQTPVRCQYHNGQSYRGFTTGTTCQSWSSMTPPHRHQRTPENYPNDGLTMYCNRN 173
Db 9 ENKDYRNCIIIGKSGYKGTSTKSGIKQCPWNSMIPHEHSLFPSSYRGKDLQENYCRN 68

Qy 174 PDADT-GPWCFTTDPSTIRWEYCNLRCSDDTEGTVVAPTVIQVPSLGLPPSQDCNFGCK 232
Db 69 PRGEGGPGWCFSTNPEVRYEVCIDIPQCSEV-----CMTCNGE 106

Qy 233 GYR 235
Db 107 STR 109

RESULT 13
Q8BS17 PRELIMINARY; PRT; 209 AA.
AC Q8BS17;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Plasminogen.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK040840; BAC30717.1; -.
SQ SEQUENCE 209 AA; 23192 MW; FA5880C13A7FA8E7 CRC64;

Query Match 14.6%; Score 260; DB 11; Length 209;
Best Local Similarity 44.3%; Pred. No. 2.6e-15;
Matches 47; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

Qy 114 EQTPVRCQYHNGQSYRGFTTGTTCQSWSSMTPPHRHQRTPENYPNDGLTMYCNRN 173
Db 95 EKRVLSECKYIGNGYRGCTWSTKSGVACQKGNATFHPVNYSPSTHNEGIGENYCRN 154

Qy 174 PDADT-GPWCFTTDPSTIRWEYCNLRCSDDTEGTVVAPTVIQVPSL 218
Db 155 PDNDEQGPWCFTTDPKRYDYCNIPCEGRGGS-RKPPSVCLSPGI 199

RESULT 14
Q9BYM0 PRELIMINARY; PRT; 208 AA.
ID Q9BYM0
AC Q9BYM0
DT 01-JUN-2001 (TRENBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hepatocyte growth factor (Fragment).
GN HGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91369928; PubMed=1832556;
RA Miyazawa K., Kitamura A., Kitamura N.;
RT "Structural organization and the transcription initiation site of the
  human hepatocyte growth factor gene."
RL Biochemistry 30:9170-9176(1991).
CC 1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; M75971; AAG53459.1; -.
DR EMBL; M75967; AAG53459.1; JOINED.
DR EMBL; M75966; AAG53459.1; JOINED.
DR EMBL; M75968; AAG53459.1; JOINED.
DR EMBL; M75969; AAG53459.1; JOINED.
DR HSP; P14210; 1BHT.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; PAN_app.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00024; PAN; 1.
DR PRINTS; PR00018; KRINGLE
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle
FT NON_TER 208
SQ SEQUENCE 208 AA; 23931 MW; AE9C50DE5A86B37B CRC64;

Query Match 14.0%; Score 249.5; DB 4; Length 208;
Best Local Similarity 47.2%; Pred. No. 2.2e-14;
Matches 42; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

Qy 114 EQTPVRCQYHNGQSYRGFTTGTTCQSWSSMTPPHRHQRTPENYPNDGLTMYCNRN 173
Db 120 ENKDYRNCIIIGKSGYKGTSTKSGIKQCPWNSMIPHEHSLFPSSYRGKDLQENYCRN 179

Qy 174 PDADT-GPWCFTTDPSTIRWEYCNLRCSDD 201
Db 180 PRGEGGPGWCFSTNPEVRYEVCIDIPQCSE 208

RESULT 15
Q13494 PRELIMINARY; PRT; 210 AA.
ID Q13494
AC Q13494;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE HGF agonist/antagonist.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96278713; PubMed=8662798;
RA Ciocce V., Csaky K.G., Chan A.M.L., Bottaro D.P., Taylor W.G.,
  Jensen R., Aaronson S.A., Rubin J.S.;
RT "Hepatocyte growth factor (HGF)/NK1 is a naturally occurring
  HGF/scatter factor variant with partial agonist/antagonist activity."
RL J. Biol. Chem. 271:13110-13115(1996).
CC 1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

```

```

DR EMBL; U46010; AAC50539.1; -.
DR HSP; P14210; IBHT.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR Pfam; PF00051; kringle; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 210 AA; 24116 MW; 94A6EE9C50DE5A86 CRC64;

Query Match 14.0%; Score 249.5; DB 4; Length 210;
Best Local Similarity 47.2%; Pred. No. 2.3e-14;
Matches 42; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 114 EQTPVVRQCYHGNGSQYRGTFSTVTGRTCCQSWSSMTFRRHQRTPENYPNDGLTMYNCRN 173
Db | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
120 ENKDYIRNCILGKRGYKGTVSITKSGIKQCPWSSMIPHERSFLPSSYRGKDLQENYCRN 179

QY 174 PDADT-GPWCFTTDFSIWEYCNLTRGSD 201
Db | : | | | | : | | : | | : | | : | | : | | : | | : | | : | |
180 PRGEGGPGWCFTSNPEVRYEVCIDIPQCSE 208

```

Search completed: September 29, 2003, 08:01:57
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:03:47 ; Search time 82 Seconds
(without alignments)
176.148 Million cell updates/sec

Title: US-10-088-548-4
Perfect score: 539
Sequence: 1 KSPVDDCYHGDFRSYRGIS.....DPCVRWEYCNLTQCSSETS 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 697866

```
Minimum DB seq length: 0
Maximum DB seq length: 91
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Database :
A_Geneseq_13Jun03.*
1: /SDSI/cgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SDSI/cgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SDSI/cgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SDSI/cgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SDSI/cgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SDSI/cgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SDSI/cgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SDSI/cgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SDSI/cgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SDSI/cgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SDSI/cgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SDSI/cgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SDSI/cgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SDSI/cgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SDSI/cgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SDSI/cgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SDSI/cgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SDSI/cgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SDSI/cgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SDSI/cgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SDSI/cgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SDSI/cgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SDSI/cgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SDSI/cgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Result No.	%			DB	ID	Description
	Score	Match	Length			
1	539	100.0	91	22	AAV72945	Human apolipoprote
2	415	77.0	90	21	AAB01909	Human plasminogen
3	405	75.1	84	21	AAV77720	Human plasminogen
4	402	74.6	89	22	AAV72946	Human apolipoprote
5	397	73.7	82	10	AAV93231	Plasminogen kringl
6	383	71.1	78	17	AAW07560	Human kringle 4
7	377	69.9	83	21	AAB08414	Amino acid sequenc
8	350	64.9	78	17	AAW07559	Murine kringle 4
9	339.5	63.0	86	23	ABP02801	Human ORFX protein

10	285	52.9	78	17	AAM07554
11	281	52.1	78	17	AAM07555
12	281	52.1	78	17	AAM07556
13	278	51.6	78	17	AAM07557
14	277	51.4	83	21	AA01908
15	274	50.8	83	21	AA08412
16	273	50.6	78	17	AAM07558
17	257.5	47.8	90	12	AAFI4661
18	256.5	47.6	89	20	AAV02110
19	255.5	47.4	89	20	AAV02134
20	255.5	47.4	89	20	AAV02133
21	252.5	46.8	79	9	AA081360
22	252.5	46.8	79	17	AA086221
23	251	46.6	90	21	AA01914
24	250.5	46.5	79	17	AAM07546
25	250.5	46.5	84	10	AA093230
26	250.5	46.5	84	21	AA01906
27	249.5	46.3	79	17	AAM07548
28	249.5	46.3	82	21	AAV77719
29	248	46.0	78	17	AAM07553
30	247.5	45.9	79	17	AAM07547
31	247.5	45.9	79	17	AAM07545
32	246	45.6	91	21	AAV58868
33	245.5	45.5	79	17	AAM07544
34	244	45.3	78	17	AAM07552
35	244	45.3	81	10	AA090598
36	244	45.3	85	21	AA01907
37	243.5	45.2	78	10	AA094401
38	242	44.9	78	17	AAM07550
39	241	44.7	78	17	AAM07551
40	239	44.3	78	21	AAV77721
41	237.5	44.1	84	21	AA080410
42	236	43.8	85	21	AA080413
43	232	43.0	52	22	ABG51665
44	232	43.0	52	22	AB031599
45	232	43.0	52	22	AB036320

Murine kringle 3.
Human kringle 3.
Rhesus kringle 3.
Porcine kringle 3.
Human plasminogen
Amino acid sequence
Bovine kringle 3.
kringle-1 of plasmin
A multifunctional
SEQ ID 113 of WO99
SEQ ID 113 of WO99
K1 domain of plasmin
Plasminogen kringle
Human plasminogen
Rhesus kringle 1.
Plasminogen kringle
Human plasminogen
Bovine kringle 1.
Human plasminogen
Bovine kringle 2.
Porcine kringle 1.
human kringle 1.
Human plasminogen
murine kringle 1.
Porcine kringle 2.
Synthetic N-termin
Human plasminogen
Sequence encoding
Human kringle 2.
Rhesus kringle 2.
Human plasminogen
Amino acid sequence
Human liver peptid
Peptide #4250 enco
Peptide #4326 enco

ALIGNMENTS

RESULT 1	
AAV72945	
ID AAV72945 standard; Protein; 91 AA.	
XX AC AAV72945;	
DT 13-JUN-2001 (first entry)	
XX Human apolipoprotein(a) kringle domain IV36, Lk6 protein..	
DE XX	
KW Human; angiogenesis inhibitor; LK6; apolipoprotein(a) kringle domain;	
KW angiogenesis-mediated disease; cancer; rheumatoid arthritis; therapy;	
KW cytostatic; antiarheumatic; antiarthritic; antipsoriatic; psoriasis;	
KW ocular angiogenic disease; endothelial cell proliferation; tumour;	
KW cell migration.	
XX XX	
OS Homo sapiens.	
XX XX	
PN WO200119868-A1.	
XX XX	
PD 22-MAR-2001.	
XX XX	
PF 15-SEP-1999; 99WO-KR00554.	
XX XX	
PR 15-SEP-1999; 99WO-KR00554.	
XX XX	
PA (MOGA-) MOGAM BIOTECHNOLOGY RES INST.	
XX XX	
PI Chang J, Kim JS, Park EJ, Yum J, Chung S;	
XX XX	
WI WI; 2001-244787/25.	
DR N-PSDB; AAD03256.	

XX PT Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
PT diseases, e.g. cancer and rheumatoid arthritis, has human
PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
PT sequence -
XX
XX Claim 1; Page 44-45; 50pp; English.
XX
XX The present sequence is human LK6 protein which contains the amino
CC acid sequence of human apolipoprotein(a) kringle domain IV36
CC (KIV36). The human apolipoprotein(a) kringle domains IV36 (LK6
CC protein), IV37 (LK7 protein) and V38 (LK8 protein) together form the
CC angiogenesis inhibitor, LK68 protein. LK68, LK6, LK7 and LK8 are
CC inhibitors of endothelial cell proliferation, cell migration and
CC normal development of capillaries in the chick embryo chorioallantoic
CC membrane (CAM). LK68 protein, its single kringles or their functional
CC equivalents, are useful for treating angiogenesis-mediated diseases,
CC such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic
CC disease in animals or humans. LK68 is useful as an anticancer agent and
CC also for inhibiting primary tumour growth.
XX
XX Sequence 91 AA;
SQ Query Match 100.0%; Score 539; DB 22; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.5e-42;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPVVQCYHGGRSYRISSTVTGRTCSWSSMTPHWHQRTPENYNAGLTENYCRNP 60
DB 1 KSPVVQCYHGGRSYRISSTVTGRTCSWSSMTPHWHQRTPENYNAGLTENYCRNP 60

QY 61 DSGKQPCWYTTDPCVWEYCNLTQCSSETS 91
DB 61 DSGKQPCWYTTDPCVWEYCNLTQCSSETS 91

RESULT 2
ID AAB01909
XX AAB01909 standard; Protein; 90 AA.
AC AAB01909;
XX
XX 18-SEP-2000 (first entry)
XX Human plasminogen kringle 4 (Val354-Val443).
XX
XX Plasminogen; human; kringle domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
XX Homo sapiens.
OS
XX US6057122-A.
FN
XX
XX 02-MAY-2000.
PD
XX
XX 05-MAY-1997; 97US-0851350.
PF
XX
XX 03-MAY-1996; 96US-0643219.
PR
XX 03-APR-1997; 97US-0832087.
XX
XX (ABBO) ABBOTT LAB.
PA
XX Davidson DJ;
XX
XX WPI; 2000-349573/30.
DR
XX Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX

PS Example 17; Page -; 48pp; English.
XX
XX The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract, tumours
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukaemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC in an exemplification of the invention.
CC Note: This sequence is not shown in the specification, but is derived
CC from the full length human plasminogen sequence (AAB01887) shown in
CC figure 1.
XX
XX Sequence 90 AA;
SQ Query Match 77.0%; Score 415; DB 21; Length 90;
Best Local Similarity 79.3%; Pred. No. 4.1e-31;
Matches 69; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 4 VVQDCYHGDGRSYRISSTVTGRTCSWSSMTPHWHQRTPENYNAGLTENYCRNPDSG 63
DB 1 VVQDCYHGDGRSYRISSTVTGRTCSWSSMTPHWHQRTPENYNAGLTENYCRNPDSG 60

QY 64 KQPCWYTTDPCVWEYCNLTQCSSETS 90
DB 61 KQPCWYTTDPCVWEYCNLTQCSSETS 87

RESULT 3
AAY77720
ID AAY77720 standard; protein; 84 AA.
XX
XX AAY77720;
XX
XX 12-MAY-2000 (first entry)
XX Human plasminogen kringle 4 sequence.
DE
XX
XX Kringle 4; plasminogen; anti-angiogenic; angiogenesis; angiotatin;
KW cytostatic; antiarthritic; antirheumatic; antidiabetic; ophthalmological;
KW immunosuppressant; vasotropic; vulnery; antiarteriosclerotic; human;
KW dermatological; cancer; tumour; birth control; vascularization.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Region 59..75
FT /note= "specifically claimed fragment (AAY77723)"
XX
XX WO200003726-A1.
FN
XX
XX 27-JAN-2000.
PD
XX
XX 07-JUL-1999; 99WO-US15271.
PF
XX
XX 14-JUL-1998; 98US-0092831.
PR
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Ji R, Trail PA;
XX

DR WPI; 2000-171200/15.

XX Novel lysine binding fragments angiotatin used as antiangiogenic

PT agents in the treatment of cancer, diabetic retinopathy, rheumatoid

PT arthritis, psoriasis, atherosclerotic plaque formation, and other

PT angiogenesis diseases -

XX Disclosure; Fig 3; 30pp; English.

XX The invention provides fragments of kringle 1, 2 or 4 of human

CC plasminogen that contain a lysine binding site and have anti-angiogenic

CC activity. The peptides of the invention function as antiangiogenic

CC agents, for the treatment process and diseases involving angiogenesis.

CC Such diseases include cancers such as solid tumours, blood born tumours

CC such as leukemias, tumours metastasis, benign tumours such as

CC hemangiomas, acoustic acuronas, neurofibromas, trachomas and pyogenic

CC granulomas, rheumatoid arthritis, ocular angiogenic diseases such as

CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,

CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,

CC rubecosis, Osler-Webber syndrome, myocardial angiogenesis, plaque

CC neovascularization, telangiectasia, hemophilic joints, angiofibroma,

CC and wound granulation. The fragments are also useful in treatment of

CC disease of excessive or abnormal stimulation of endothelial cells. These

CC diseases include intestinal adhesions, atherosclerosis, scleroderma, and

CC hypertrophic scars. The fragments can be used as birth control agents by

CC preventing vascularization required for embryo transplantation. The

CC present sequence represents the kringle 4 sequence of human plasminogen

CC (angiotatin).

XX Sequence 84 AA;

SQ Query Match 75.1%; Score 405; DB 21; Length 84;

Best Local Similarity 79.8%; Pred. No. 3.2e-30;

Matches 67; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 5 VDCYHGDSYRGISSTVTGTGTCQSSMIPHWHQRTPENYPNAGLTENYCRNPDGK 64

Db 1 VDCYHGDSYRGISSTVTGTGTCQSSMIPHWHQRTPENYPNAGLTENYCRNPDGK 60

Qy 65 QPWCFTTDPVRYEYCNLTQCSFT 88

Db 61 GPWCFTTDPVRYEYCNLTQCSFT 84

RESULT 4

AAY72946

ID AAY72946 standard; Protein; 89 AA.

XX AAY72946;

XX 13-JUN-2001 (first entry)

XX Human apolipoprotein(a) kringle domain IV37, LK7 protein.

XX Human; angiogenesis inhibitor; LK7; apolipoprotein(a) kringle domain;

XX angiogenesis-mediated disease; cancer; rheumatoid arthritis; therapy;

XX cytosolic; antithematic; antiarthritic; antipsoriatic; psoriasis;

XX ocular angiogenic disease; endothelial cell proliferation; tumour;

XX cell migration.

XX Homo sapiens.

XX WO200119868-A1.

XX 22-MAR-2001.

XX 15-SEP-1999; 99WO-KR00554.

XX 15-SEP-1999; 99WO-KR00554.

XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.

XX Chang J, Kim JS, Park EJ, Yum J, Chung S;

XX WPI; 2001-244787/25.

DR N-PSDB; AAD03257.

XX Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated

PT diseases, e.g. cancer and rheumatoid arthritis, has human

PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid

PT sequence -

XX Claim 2; Page 45-46; 50pp; English.

XX The present sequence is human LK7 protein which contains the amino

CC acid sequence of human apolipoprotein(a) kringle domain IV37

CC (KIV37). The human apolipoprotein(a) kringle domains IV36 (LK6

CC protein), IV37 (LK7 protein) and V38 (LK8 protein) together form the

CC angiogenesis inhibitor, LK68 protein. LK68, LK6, LK7 and LK8 are

CC inhibitors of endothelial cell proliferation, cell migration and

CC normal development of capillaries in the chick embryo chorioallantoic

CC membrane (CAM). LK68 protein, its single kringles or their functional

CC equivalents, are useful for treating angiogenesis-mediated diseases,

CC such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic

CC disease in animals or humans. LK68 is useful as an anticancer agent and

CC also for inhibiting primary tumour growth.

XX Sequence 89 AA;

SQ Query Match 74.6%; Score 402; DB 22; Length 89;

Best Local Similarity 77.6%; Pred. No. 6.5e-30;

Matches 66; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Qy 5 VDCYHGDSYRGISSTVTGTGTCQSSMIPHWHQRTPENYPNAGLTENYCRNPDGK 64

Db 1 VDCYHGDSYRGISSTVTGTGTCQSSMIPHWHQRTPENYPNAGLTENYCRNPDGK 60

Qy 65 QPWCFTTDPVRYEYCNLTQCSFT 89

Db 61 GPWCFTTDPVRYEYCNLTQCSFT 85

RESULT 5

AAP93231

ID AAP93231 standard; peptide; 82 AA.

XX AAP93231;

XX 25-MAR-2003 (updated)

XX 03-APR-1990 (first entry)

XX Plasminogen kringle 4 domain (residues 354-435).

XX Plasminogen; activator; t-PA; fibrin; kringle domain;

XX scu-PA;

XX Homo sapiens.

XX WO8910401-A.

XX 02-NOV-1989.

XX 23-MAR-1989; 89WO-US01255.

XX 22-APR-1988; 88US-0184823.

XX (COLB) COLLABORATIVE RES INC.

XX Mao JI, Abercrombie DM;

XX WPI; 1989-339965/46.

XX N-PSDB; AAN92238.

XX Modified plasminogen activator - having greater fibrin

PT selectivity and circulating half-life.

XX

PS Disclosure; fig 3b; 80pp; English.

XX The sequence is encoded by a synthetic oligonucleotide and is the kringle

CC 4 domain of plasminogen. It is used to construct a modified plasminogen

CC activator (mPA) of greater fibrin selectivity than normal PA. The mPA is

CC used in the treatment of pulmonary embolism, thrombosis, myocardial

CC infarct and strokes.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 82 AA;

SQ

Query Match 73.7%; Score 397; DB 10; Length 82;

Best Local Similarity 79.3%; Pred. No. 1.7e-29;

Matches 65; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 4 VVQCYHGDSRGISSTVTGRTCCQSSSMIPHHQRTPENYPNAGLTENYCRNPDSG 63

Db 1 VVQCYHNGQSTRGISSTVTGRTCCQSSSMTPRHQRTPENYPNAGLTENYCRNPDD 60

QY 64 KQPCWYTTDPCVWEYCNLTQC 85

Db 61 KGPWCFTTDPVWEYCNLKKC 82

RESULT 6

AAW07560

ID AAW07560 standard; protein; 78 AA.

XX

AC AAW07560;

XX

DT 22-JUN-1997 (first entry)

XX

DE Human kringle 4.

XX

KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;

KW macular degeneration; diabetic retinopathy.

XX

OS Homo sapiens.

XX

PN WO9635774-A2.

XX

PD 14-NOV-1996.

XX

PF 26-APR-1996; 96WO-US05856.

XX

PR 08-MAR-1996; 96US-0612788.

XX

PR 26-APR-1995; 95US-0429743.

XX

PR 22-FEB-1996; 96US-0605598.

XX

PA (CHIL-) CHILDRENS MEDICAL CENT.

XX

PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;

XX

PI WPI; 1996-518662/51.

XX

DR Use of angiotatin fragments or aggregates - for inhibiting

XX endothelial cell proliferation and treating angiogenesis-mediated

PT diseases, e.g. cancer, arthritis or diabetic retinopathy

PT

PS Claim 4; Page 116; 203pp; English.

XX

XX The invention relates to new methods and compositions for

CC inhibiting endothelial cell proliferation, using as active component

CC an angiotatin fragment, a combination of angiotatin fragments, or

CC aggregate angiotatin. The fragment is preferably derived from murine,

CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,

CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle

CC 1-4 or kringle 1-4BKLs protein. The aggregate angiotatin has a Mol. Wt.

CC of 45-65 kD and is derived from a plasminogen fragment beginning at

CC approximately amino acid number 98 of murine, human, Rhesus, porcine or

CC bovine plasminogen. The active component can be used for treating

CC angiogenesis-mediated diseases such as cancer, arthritis, macular

CC degeneration and diabetic retinopathy. It can also be used to develop

CC antibodies for use in diagnosis, detection and therapy.

CC The present sequence, human kringle 4, is a specific angiotatin

CC fragment which can be used in the invention.

XX

SQ Sequence 78 AA;

Query Match 71.1%; Score 383; DB 17; Length 78;

Best Local Similarity 79.5%; Pred. No. 3.2e-28;

Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 8 CYHGDGSRGYSSTVTGRTCCQSSSMIPHHQRTPENYPNAGLTENYCRNPDSGKQPW 67

Db 1 CYHGDGQSYRGTSSTTTGKQSSSMTPRHQKTPENYPNAGLTENYCRNPADKGPW 60

QY 68 CYTTDPCVWEYCNLTQC 85

Db 61 CFTTDPVWEYCNLKKC 78

RESULT 7

AAW08414

ID AAW08414 standard; Protein; 83 AA.

XX

AC AAW08414;

XX

DT 20-DEC-2000 (first entry)

XX

DE Amino acid sequence of kringle 4 of human plasminogen.

XX

KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;

KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;

KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;

KW adult respiratory distress syndrome; Castelman's disease; psoriasis;

KW hepatitis; aneurysm; renal disease; haemangioma.

XX

OS Homo sapiens.

XX

PN WO2000049871-A1.

XX

PD 31-AUG-2000.

XX

PF 24-FEB-2000; 2000WO-US04798.

XX

PR 24-FEB-1999; 99US-0121341.

XX

PR 25-FEB-1999; 99US-0121633.

XX

PR 18-NOV-1999; 99US-0166176.

XX

PA (FORD-) FORD HEALTH SYSTEM HENRY.

XX

PI Dou D, Chopp M, Wang L, Mikkelsen T;

XX

PI WPI; 2000-572016/53.

XX

DR Use of kringle protein and kringle derived from plasminogen and

XX composition comprising kringle proteins for treating tumor and

PT atherosclerosis, arthritis and retinopathy -

PT

PS Disclosure; Fig 6; 163pp; English.

XX

XX The specification describes a human polypeptide which is a potent

CC angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle

CC proteins, or a kringle derived from human tissue plasminogen activator

CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED

CC or tPA are useful for treating tumours, as well as atherosclerosis,

CC arthritis, retinopathy and other similar diseases. KED is also useful

CC for the treatment of diseases such as bronchial vascular congestion,

CC inflammatory bowel disease, adult respiratory distress syndrome,

CC Castelman's disease, psoriasis, hepatitis, aneurysm, renal disease

CC and haemangioma. The present sequence represents kringle 4 of human

CC plasminogen, which is used in the course of the invention.

XX

SQ Sequence 83 AA;

Query Match 69.9%; Score 377; DB 21; Length 83;
Best Local Similarity 76.8%; Pred. No. 1.2e-27;
Matches 63; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 5 VDCYHGDSYRGISSTVTGRTCSWSSMIPHHQRTPTENYNAGLTENYCRNPDGSK 64
DB 1 VQCYHGDGQSYRGISSTVTGRTCSWSSMTPPHHQKTPENYNGLTWNYCRNPDAK 60

QY 65 QPCYTTDPCVRWEYCNLTQCS 86
DB 61 GPTCTDPSVRWEYCNLKCS 82

RESULT 8
AAW07559
ID AAW07559 standard; protein; 78 AA.

XX AC AAW07559;

XX XX 22-JUN-1997 (first entry)

XX DE Murine kringle 4.

XX KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
XX KW macular degeneration; diabetic retinopathy.

XX OS Mus musculus.

XX PN WO9635774-A2.

XX PD 14-NOV-1996.

XX PF 26-APR-1996; 96WO-US05856.

XX PR 08-MAR-1996; 96US-0612788.

XX PR 26-APR-1995; 95US-0429743.

XX PR 22-FEB-1996; 96US-0605598.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;

XX DR WPI; 1996-518662/51.

XX PT Use of angiotatin fragments or aggregates - for inhibiting

XX PS Claim 4; Page 115; 203pp; English.

XX CC The invention relates to new methods and compositions for
XX CC inhibiting endothelial cell proliferation, using as active component
XX CC an angiotatin fragment, a combination of angiotatin fragments, or
XX CC aggregate angiotatin. The fragment is preferably derived from murine,
XX CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
XX CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
XX CC 1-4 or kringle 1-4BKIS protein. The aggregate angiotatin has a Mol. wt.
XX CC of 45-65 kD and is derived from a plasminogen fragment beginning at
XX CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
XX CC bovine plasminogen. The active component can be used for treating
XX CC angiogenesis-mediated diseases such as cancer, arthritis, macular
XX CC degeneration and diabetic retinopathy. It can also be used to develop
XX CC antibodies for use in diagnosis, detection and therapy.
XX CC The present sequence, murine kringle 4, is a specific angiotatin
XX CC fragment which can be used in the invention, and represents amino
XX CC acids 377-454 of murine plasminogen.

XX SQ Sequence 78 AA;

Query Match 64.9%; Score 350; DB 17; Length 78;
Best Local Similarity 70.5%; Pred. No. 3.6e-25;
Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 CYHGDSYRGISSTVTGRTCSWSSMIPHHQRTPTENYNAGLTENYCRNPDGSKQFW 67
DB 1 CQSDGQSGRTGTSSTVTGRTCSWSSMTPPHHQKTPENYNGLTWNYCRNPDKGFW 60

QY 68 CYTTDPCVRWEYCNLTQCS 85
DB 61 CYTTDPSVRWEYCNLKRC 78

RESULT 9
ABP02801
ID ABP02801 standard; Protein; 86 AA.

XX AC ABP02801;

XX XX 24-JUN-2002 (first entry)

XX DE Human OREF protein sequence SEQ ID NO:5584.

XX KW Human; open reading frame; OREF; gene therapy; cancer; cirrhosis;
XX KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX KW hypertension; hypothyroidism; cholesterol ester storage disease;
XX KW immune deficiency; immune disorder; infectious disease;
XX KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX KW myasthenia gravis.

XX OS Homo sapiens.

XX PN WO200192523-A2.

XX PD 06-DEC-2001.

XX PF 29-MAY-2001; 2001WO-US10836.

XX PR 30-MAY-2000; 2000US-206132P.

XX PR 29-AUG-2000; 2000US-228716P.

XX PA (CURA-) CURAGEN CORP.

XX PI Shinkets RA, Leach MD;

XX DR WPI; 2002-106308/14.

XX DR N-PSDB; ABN18553.

XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
XX PT preventing and treating cardiovascular disease, neurodegenerative,
XX PT hyperproliferative disorders and autoimmune disorders

XX PS Disclosure; SEQ ID 5584; 1037pp; English.

XX CC The present invention describes substantially purified human proteins
XX CC (referred to as open reading frame, OREF, where X is 1-11491 (see Table 1
XX CC in the specification). ABN15762 to ABN27252 encode the human OREF
XX CC proteins given in ABP00010 to ABP11500. OREF proteins are useful for
XX CC treating or preventing a pathology associated with an OREF-associated
XX CC disorder in humans, and in the manufacture of a medicament for treating a
XX CC syndrome associated with OREF-associated disorder. OREF polynucleotide
XX CC sequences can be used in gene therapy. OREF sequences can be used in the
XX CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
XX CC osteoarthritis, neurodegenerative disorders, disorders related to organ
XX CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
XX CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
XX CC storage disease, various immune deficiencies and disorders, infectious
XX CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
XX CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
XX CC disease and autoimmune inflammatory eye disease. OREF proteins are also
XX CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
XX CC bone degenerative disorders, or periodontal disease, and for gut
XX CC protection or regeneration and treatment of lung or liver fibrosis,
XX CC reperfusion injury in various tissues and conditions resulting from

CC The invention relates to new methods and compositions for
CC inhibiting endothelial cell proliferation, using as active component
CC an angiotensin fragment, a combination of angiotensin fragments, or
CC aggregate angiotensin. The fragment is preferably derived from murine,
CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotensin has a Mol. Wt.
CC of 45-65 kD and is derived from a plasminogen fragment beginning at
CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
CC bovine plasminogen. The active component can be used for treating
CC angiotensin-mediated diseases such as cancer, arthritis, macular
CC degeneration and diabetic retinopathy. It can also be used to develop
CC antibodies for use in diagnosis, detection and therapy.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:18:25 ; Search time 29 Seconds
(without alignments)
132.769 Million cell updates/sec

Title: US-10-088-548-4
Perfect score: 539
Sequence: 1 KSPVYQDCYHGDGRSYRGIS.....DPCYRWEYCNLTQCSFTSG 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 224404

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2.6/ptodata/1/iaa/PTQUS_COMB.pep.*
6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	75.1	84	4	US-09-348-953-2
2	383	71.1	78	2	US-08-612-788-23
3	383	71.1	78	2	US-08-612-788-23
4	383	71.1	78	4	US-09-066-028-23
5	383	71.1	78	4	US-09-335-325-23
6	350	64.9	78	2	US-08-612-788-22
7	350	64.9	78	3	US-09-066-028-22
8	350	64.9	78	4	US-09-335-325-22
9	285	52.9	78	2	US-08-612-788-17
10	285	52.9	78	3	US-09-066-028-17
11	285	52.9	78	4	US-09-335-325-17
12	281	52.1	78	2	US-08-612-788-18
13	281	52.1	78	2	US-08-612-788-19
14	281	52.1	78	3	US-09-066-028-18
15	281	52.1	78	3	US-09-066-028-19
16	281	52.1	78	4	US-09-335-325-18
17	281	52.1	78	4	US-09-335-325-19
18	278	51.6	78	2	US-08-612-788-20
19	278	51.6	78	3	US-09-066-028-20
20	278	51.6	78	4	US-09-335-325-20
21	277	51.4	78	2	US-08-612-788-4
22	273	50.6	78	2	US-08-612-788-21
23	273	50.6	78	3	US-09-066-028-21
24	273	50.6	78	4	US-09-335-325-21
25	251	46.6	90	1	US-08-643-219-4
26	251	46.6	90	2	US-09-131-995-4
27	251	46.6	90	2	US-08-832-087B-4

28	251	46.6	90	3	US-09-132-154-4	Sequence 4, Appli
29	250.5	46.5	79	2	US-08-612-788-9	Sequence 9, Appli
30	250.5	46.5	79	3	US-09-066-028-9	Sequence 9, Appli
31	250.5	46.5	79	4	US-09-335-325-9	Sequence 9, Appli
32	249.5	46.3	79	2	US-08-612-788-11	Sequence 11, Appli
33	249.5	46.3	79	3	US-09-066-028-11	Sequence 11, Appli
34	249.5	46.3	79	4	US-09-335-325-11	Sequence 11, Appli
35	249.5	46.3	82	4	US-09-348-953-1	Sequence 1, Appli
36	248	46.0	78	2	US-08-612-788-16	Sequence 16, Appli
37	248	46.0	78	3	US-09-066-028-16	Sequence 16, Appli
38	248	46.0	78	4	US-09-335-325-16	Sequence 16, Appli
39	247.5	45.9	79	2	US-08-612-788-8	Sequence 8, Appli
40	247.5	45.9	79	2	US-08-612-788-10	Sequence 10, Appli
41	247.5	45.9	79	2	US-08-763-528A-2	Sequence 2, Appli
42	247.5	45.9	79	3	US-09-066-028-8	Sequence 8, Appli
43	247.5	45.9	79	3	US-09-066-028-10	Sequence 10, Appli
44	247.5	45.9	79	4	US-09-335-325-8	Sequence 8, Appli
45	247.5	45.9	79	4	US-09-335-325-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-09-348-953-2
; Sequence 2, Application US/09348953
; Patent No. 6538103
; GENERAL INFORMATION:
; APPLICANT: JI, Richard Wei-Dong
; APPLICANT: Trail, Pamela A.
; TITLE OF INVENTION: LYSINE BINDING FRAGMENTS OF ANGIOSTATIN
; FILE REFERENCE: D311Sequences
; CURRENT APPLICATION NUMBER: US/09/348,953
; CURRENT FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/052,831
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 84
; TYPE: PRT
; ORGANISM: human
US-09-348-953-2

Query Match 75.1%; Score 405; DB 4; Length 84;
Best Local Similarity 79.8%; Pred. No. 1.3e-36;
Matches 67; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
QY 5 VDCYHGDGRSYRGISSTVTGTCTQSSSMIPHWQHTFENYFNAGLTENYCRNPDGK 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 1 VDCYHGDGRSYRGISSTVTGTCTQSSSMIPHWQHTFENYFNAGLTENYCRNPDGK 60
QY 65 QPWCYTTDPCYRWEYCNLTQCSFT 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 61 GPWCFTTDPYRWEYCNLTQCSGT 84
RESULT 2
US-08-612-788-23
; Sequence 23, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia


```
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K4
/ US-09-066-028-23
```

```
Query Match 71.1%; Score 383; DB 3; Length 78;
Best Local Similarity 79.5%; Pred. No. 2.9e-34;
Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 8 CYHGDGRSYRGISSTVTGRTCSQSSMIPHPHWHQRTPEYNPNAGLTENYCRNPDSCQKQPW 67
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CYHGDQSYRGTSSTTTGKKCSQSSMTPPHRHQKTPENYPNAGLTENYCRNPDADKQGPW 60
```

```
QY 68 CYTTPDCVRYEYCNLTQC 85
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CFTTDPVSRWEYCNLKKC 78
```

```
RESULT 5
US-09-335-325-23
; Sequence 23, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
```

```
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
```

```
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K4
/ SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-335-325-23
```

```
Query Match 71.1%; Score 383; DB 4; Length 78;
Best Local Similarity 79.5%; Pred. No. 2.9e-34;
Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 8 CYHGDGRSYRGISSTVTGRTCSQSSMIPHPHWHQRTPEYNPNAGLTENYCRNPDSCQKQPW 67
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CYHGDQSYRGTSSTTTGKKCSQSSMTPPHRHQKTPENYPNAGLTENYCRNPDADKQGPW 60
```

```
QY 68 CYTTPDCVRYEYCNLTQC 85
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CFTTDPVSRWEYCNLKKC 78
```

```
RESULT 6
US-08-612-788-22
; Sequence 22, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; US-08-612-788-22
```


RESULT 9
US-08-612-788-17
; Sequence 17, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
; US-08-612-788-17
Query Match 52.9%; Score 285; DB 2; Length 78;
Best Local Similarity 59.0%; Pred. No. 1e-23;
Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;
QY 8 CYHGDGRYRGISSTVTGTCQSWSSMIPHWQRTPENYPNAGLTENYCRNPDGKOPW 67
Db 1 CLKRGENYRGTVSVTGTCQSWSSMIPHWQRTPENYPNAGLTENYCRNPDGKOPW 60
QY 68 CYTTDPQVRYEYCNLTQC 85
Db 61 CYTTDSQLRWYCEIPSC 78
RESULT 10
US-09-066-028-17
; Sequence 17, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use

; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
; US-09-066-028-17
Query Match 52.9%; Score 285; DB 3; Length 78;
Best Local Similarity 59.0%; Pred. No. 1e-23;
Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;
QY 8 CYHGDGRYRGISSTVTGTCQSWSSMIPHWQRTPENYPNAGLTENYCRNPDGKOPW 67
Db 1 CLKRGENYRGTVSVTGTCQSWSSMIPHWQRTPENYPNAGLTENYCRNPDGKOPW 60
QY 68 CYTTDPQVRYEYCNLTQC 85
Db 61 CYTTDSQLRWYCEIPSC 78
RESULT 11
US-09-335-325-17
; Sequence 17, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612.788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K3
; US-08-612-788-18

Query Watch                    52.1%; Score 281; DB 2; Length 78;
Best Local Similarity          56.4%; Pred.No.2.7e-23;
Matches    44; Conservative   10; Mismatches    24; Indels      0; Gaps     0;

QY       8 CYHGDGRSYRGTSVTVTGTCQSWSMHPHQRTPEPNYNAGLTENYCNPDSGKQFW        67
           | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db       1 CLKGTGENGVNATVSHTCQHRSQTTHINRPFNFTSKNLDENYCNPDGSRAPW         60

QY       68 CYYTDPVCRWEXNYCLTQC      85
           |:||: ||||||: |
Db       61 CHTNISQVRWEYCKIPSC      78

RESULT 13
US-08-612-788-19
; Sequence 19, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiotstatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612.788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
```



```
;
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 78 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rhesus monkey
; IMMEDIATE SOURCE:
; CLONE: K3
;
US-08-612-788-19
Query Match 52.1%; Score 281; DB 2; Length 78;
Best Local Similarity 57.7%; Pred. No. 2.7e-23;
Matches 45; Conservative 7; Mismatches 26; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGTCQSSMIPHWORHPENYPNAGLTENYCRNPDSGKQPW 67
DB 1 CLKGTGENYRGVAVTVSGHTCGWSAQTPHTHRTPEFPCKNLNDENYCRNPDSGKAPW 60
QY 68 CYTTPDCVRWEYCNLTQC 85
DB 61 CHTINSQVRWEYCKIPSC 78

RESULT 14
US-09-066-028-18
; Sequence 18, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rhesus monkey
; IMMEDIATE SOURCE:

;
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K3
;
US-09-066-028-18
Query Match 52.1%; Score 281; DB 3; Length 78;
Best Local Similarity 56.4%; Pred. No. 2.7e-23;
Matches 44; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGTCQSSMIPHWORHPENYPNAGLTENYCRNPDSGKQPW 67
DB 1 CLKGTGENYRGVAVTVSGHTCGWSAQTPHTHRTPEFPCKNLNDENYCRNPDSGKAPW 60
QY 68 CYTTPDCVRWEYCNLTQC 85
DB 61 CHTINSQVRWEYCKIPSC 78

RESULT 15
US-09-066-028-19
; Sequence 19, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rhesus monkey
; IMMEDIATE SOURCE:
```

CLONE: K3
US-09-066-028-19

Query Match 52.1%; Score 281; DB 3; Length 78;
Best Local Similarity 57.7%; Pred. No. 2.7e-23;
Matches 45; Conservative 7; Mismatches 26; Indels 0; Gaps 0;
QY 8 CYEGDGRSYRGISSTTVTGRTCQSSSMIPHWHTPENYPNAGLTENYCRNPDGKQPW 67
Db 1 CLKGTGENYRGDVAYTVSGHTCHGWSAQTPHTHTNTPENFPCKNLDENYCRNPDGEKAPW 60
QY 68 CYTDPQVRWEYCNLTQC 85
Db 61 CYTNSQVRWEYCKIPSC 78

Search completed: September 29, 2003, 08:23:21
Job time : 31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:22:08 : Search time 63 Seconds

(without alignments)
218.555 Million cell updates/sec

Title: US-10-088-548-4

Perfect score: 539

Sequence: 1 KSPVQDCYHGDRSYRGIS.....DPCVRWEYCNLTQCSIESG 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 203300

Minimum DB seq length: 0

Maximum DB seq length: 91

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:**

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	383	71.1	78	9	US-09-753-064-5
2	383	71.1	78	9	US-09-761-120-23
3	383	71.1	78	10	US-09-335-325-23
4	383	71.1	78	12	US-10-267-137-9
5	383	71.1	78	15	US-10-131-241-23
6	350	64.9	78	9	US-09-761-120-22
7	350	64.9	78	10	US-09-335-325-22
8	350	64.9	78	15	US-10-131-241-22
9	285	52.9	78	9	US-09-761-120-17
10	285	52.9	78	10	US-09-335-325-17
11	285	52.9	78	15	US-10-131-241-17
12	281	52.1	78	9	US-09-761-120-19
13	281	52.1	78	10	US-09-335-325-18
14	281	52.1	78	10	US-09-335-325-19
15	281	52.1	78	15	US-10-131-241-18

```
16 281 52.1 78 15 US-10-131-241-19
17 278 51.6 78 9 US-09-761-120-20
18 278 51.6 78 10 US-09-335-325-20
19 278 51.6 78 15 US-10-131-241-20
20 277 51.4 78 9 US-09-753-064-4
21 277 51.4 78 9 US-09-761-120-18
22 273 50.6 78 9 US-09-761-120-21
23 273 50.6 78 10 US-09-335-325-21
24 273 50.6 78 15 US-10-131-241-21
25 271 50.3 78 12 US-10-267-137-12
26 250.5 46.5 79 9 US-09-761-120-9
27 250.5 46.5 79 10 US-09-335-325-9
28 250.5 46.5 79 15 US-10-131-241-9
29 249.5 46.3 79 9 US-09-761-120-11
30 249.5 46.3 79 10 US-09-335-325-11
31 249.5 46.3 79 15 US-10-131-241-11
32 248 46.0 78 9 US-09-761-120-16
33 248 46.0 78 10 US-09-335-325-16
34 248 46.0 78 15 US-10-131-241-16
35 247.5 45.9 79 9 US-09-753-064-2
36 247.5 45.9 79 9 US-09-761-120-8
37 247.5 45.9 79 9 US-09-761-120-10
38 247.5 45.9 79 10 US-09-335-325-8
39 247.5 45.9 79 10 US-09-335-325-10
40 247.5 45.9 79 12 US-10-267-137-7
41 247.5 45.9 79 15 US-10-131-241-8
42 247.5 45.9 79 15 US-10-131-241-10
43 245.5 45.5 79 10 US-09-335-325-7
44 245.5 45.5 79 15 US-10-131-241-7
45 244 45.3 78 9 US-09-761-120-15
```

ALIGNMENTS

RESULT 1

```
US-09-753-064-5
: Sequence 5, Application US/09753064
: Patent No. US200100166441
: GENERAL INFORMATION:
: APPLICANT: Cao, Yihai
: Folkman, M. Judah
: TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
: and Method of Use
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Jones & Askew, LLP
: STREET: 191 Peachtree Street, 37th Floor
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: US
: ZIP: 30303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/753,064
: FILING DATE: 29-Dec-2000
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/763,528
: FILING DATE: 12-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Warren, William L.
: REGISTRATION NUMBER: 36,714
: REFERENCE/DOCKET NUMBER: 05940-0251
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-818-3700
: TELEFAX: 404-818-3799
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
```

```
Sequence 19, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 4, Appli
Sequence 18, Appli
Sequence 21, Appl
Sequence 21, Appl
Sequence 21, Appl
Sequence 12, Appl
Sequence 9, Appli
Sequence 9, Appli
Sequence 9, Appli
Sequence 11, Appl
Sequence 11, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 2, Appli
Sequence 8, Appli
Sequence 10, Appl
Sequence 10, Appl
Sequence 7, Appli
Sequence 8, Appli
Sequence 10, Appl
Sequence 7, Appli
Sequence 15, Appl
```

```

1 O'Reilly, Micheal
2 Cao, Yihai
3 Sim, B. Kim Lee
4
5 TITLE OF INVENTION: Angiotatin Fragments and Method of Use
6 NUMBER OF SEQUENCES: 45
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Jones & Askew
9 STREET: 191 Peachtree Street, 37th Floor
10 CITY: Atlanta
11 STATE: Georgia
12 COUNTRY: U.S.
13 ZIP: 30303-1769
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.30
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/09/335,325
22 FILING DATE: 17-Jun-1999
23 CLASSIFICATION: <Unknown>
24
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US/08/612,788
27 FILING DATE: <Unknown>
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Warren, William L.
30 REGISTRATION NUMBER: 36,714
31 REFERENCE/DOCKET NUMBER: 05213-0126
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: 404-818-3700
34 TELEFAX: 404-818-3799
35 INFORMATION FOR SEQ ID NO: 23:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 78 amino acids
38 TYPE: amino acid
39 STRANDEDNESS: <Unknown>
40 TOPOLOGY: linear
41 MOLECULE TYPE: protein
42 HYPOTHETICAL: NO
43 ANTI-SENSE: NO
44 FRAGMENT TYPE: N-terminal
45 ORIGINAL SOURCE:
46 ORGANISM: Homo sapiens
47 IMMEDIATE SOURCE:
48 CLONE: K4
49
50 SEQUENCE DESCRIPTION: SEQ ID NO: 23:
51
52 US-09-335-325-23
53
54 Query Match 71.1%; Score 383; DB 10; Length 78;
55 Best Local Similarity 79.5%; Pred. No. 3.8e-34;
56 Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
57
58 QY 8 CYHGDGRSYRGISSTVTGRTQSWSMIPHHWQRTPENYNAGLTENYCRNPDSGKQPW 67
59 Db 1 CYHGDGQSYRGSSVTITGKKQCSWSMTPRHQKTPENYNAGLTENYCRNPDSKGPW 60
60
61 QY 68 CYTTDFCVRWYCNLTQC 85
62 Db 61 CFTTDFSVRWYCNLKKC 78
63
64 RESULT 4
65 US-10-267-137-9
66 Sequence 9, Application US/10267137
67 Publication No. US20030148950A1
68 GENERAL INFORMATION:
69 APPLICANT: Xin, Li
70 APPLICANT: Li, Zai-Ping
71 APPLICANT: Gan, Ren-bao
72 APPLICANT: Zhou, Qing-wei
73 APPLICANT: Xu, Ren
74
75 TITLE OF INVENTION: KRINGLE DOMAIN 1 OF HUMAN HEPATOCYTE
76 GROWTH FACTOR AND USES THEREFOR
77

```

```
; FILE REFERENCE: 524282000100
; CURRENT APPLICATION NUMBER: US/10/267,137
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/328,329
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence alignment of kringle domains of
; OTHER INFORMATION: plasminogen and HGF
US-10-267-137-9

Query Match          71.1%; Score 383; DB 12; Length 78;
Best Local Similarity 79.5%; Pred. No. 3.8e-34;
Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHHQRTPENYPNAGLTENYCRNPDGKQPW 67
   |||||:||||| ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
Db 1 CYHGDGQSYRGTSSTTTGKCCQSWSSMTPHRRQKTPENYPNAGLTWNYCRNPDADKGPW 60

QY 68 CYTTDPCVWEYCNLTQC 85
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CFTTDPVWEYCNLKC 78

RESULT 5
US-10-131-241-23
; Sequence 23, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-23

Query Match          71.1%; Score 383; DB 15; Length 78;
Best Local Similarity 79.5%; Pred. No. 3.8e-34;
Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHHQRTPENYPNAGLTENYCRNPDGKQPW 67
   |||||:||||| ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
Db 1 CYHGDGQSYRGTSSTTTGKCCQSWSSMTPHRRQKTPENYPNAGLTWNYCRNPDADKGPW 60

QY 68 CYTTDPCVWEYCNLTQC 85
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CFTTDPVWEYCNLKC 78

RESULT 6
US-09-761-120-22
; Sequence 22, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasm
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Kringle 4
US-09-761-120-22

Query Match          64.9%; Score 350; DB 9; Length 78;
Best Local Similarity 70.5%; Pred. No. 1.4e-30;
Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHHQRTPENYPNAGLTENYCRNPDGKQPW 67
   |||||:||||| ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
Db 1 CYHGDGQSYRGTSSTTTGKCCQSWAAFPHRSKTPENFPDAGLEMMYCRNPDGKGPW 60

QY 68 CYTTDPCVWEYCNLTQC 85
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CFTTDPVWEYCNLKC 78

RESULT 7
US-09-335-325-22
; Sequence 22, Application US/09335325
; Patent No. US2002016471A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-335-325-22

Query Match      64.9%; Score 350; DB 10; Length 78;
Best Local Similarity 70.5%; Pred. No. 1.4e-30;
Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPEPNYVAGLTENYCRNPDSGKQPW 67
Db 1 CYQSDGOSYRGTSSTITGKKQSWAAMPFRRHSKTPENFPDAGLEMYCRNPDPGDKGPW 60

QY 68 CYTTDPCVRWEYCNLTQC 85
Db 61 CYTTDPSVRWEYCNLKRC 78

RESULT 8
US-10-131-241-22
; Sequence 22, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murinae sp.
US-10-131-241-22

Query Match      54.9%; Score 350; DB 15; Length 78;
Best Local Similarity 70.5%; Pred. No. 1.4e-30;
Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPEPNYVAGLTENYCRNPDSGKQPW 67
Db 1 CYQSDGOSYRGTSSTITGKKQSWAAMPFRRHSKTPENFPDAGLEMYCRNPDPGDKGPW 60

QY 68 CYTTDPCVRWEYCNLTQC 85
Db 61 CYTTDPSVRWEYCNLKRC 78

RESULT 9
US-09-761-120-17
; Sequence 17, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
```

```
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 3
US-09-761-120-17

Query Match      52.9%; Score 285; DB 9; Length 78;
Best Local Similarity 59.0%; Pred. No. 1.5e-23;
Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPEPNYVAGLTENYCRNPDSGKQPW 67
Db 1 CLKRGENTRGTVSVTVSGKTCQWSEQTPHRHRTPEPNFCKNLENYCRNPDGETAPW 60

QY 68 CYTTDPCVRWEYCNLTQC 85
Db 61 CYTTDQQLRWEYCEIPSC 78

RESULT 10
US-09-335-325-17
; Sequence 17, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
```

```

;
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-335-325-17

Query Match      52.98; Score 285; DB 10; Length 78;
Best Local Similarity 59.08; Pred. No. 1.5e-23;
Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

Qy      8 CYHGDGRSYRGISSTVTGTCQSSSMIPHHQRTPEYNPAGLTENYCRNPDGSKQPW 67
Db      1 CLKGRGENYRGTVSVTSGKTCQWSEQTPHRRNRTPEYNPCKNLEENYCRNPDGETAPW 60

Qy      68 CYTTPDCVRWEYCNLTQC 85
Db      61 CYTDSQLRWEYCEIPSC 78

RESULT 11
US-10-131-241-17
; Sequence 17, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murine
US-10-131-241-17

Query Match      52.98; Score 285; DB 15; Length 78;
Best Local Similarity 59.08; Pred. No. 1.5e-23;
Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

Qy      8 CYHGDGRSYRGISSTVTGTCQSSSMIPHHQRTPEYNPAGLTENYCRNPDGSKQPW 67
Db      1 CLKGRGENYRGTVSVTSGKTCQWSEQTPHRRNRTPEYNPCKNLEENYCRNPDGETAPW 60

Qy      68 CYTTPDCVRWEYCNLTQC 85
Db      61 CYTDSQLRWEYCEIPSC 78

RESULT 12
US-09-761-120-19
; Sequence 19, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmi
; FILE REFERENCE: 05940-0151 (43171-252058)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Macaca sp.
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 3
US-09-761-120-19

Query Match      52.1%; Score 281; DB 9; Length 78;
Best Local Similarity 57.7%; Pred. No. 4e-23;
Matches 45; Conservative 7; Mismatches 26; Indels 0; Gaps 0;

Qy      8 CYHGDGRSYRGISSTVTGTCQSSSMIPHHQRTPEYNPAGLTENYCRNPDGSKQPW 67
Db      1 CLKGTGENYRGDVAVTGVTGCHGWSAQTPTHNRTPENFPCKNLDENYCRNPDGEKAPW 60

Qy      68 CYTTPDCVRWEYCNLTQC 85
Db      61 CYTTSQVRWEYCKIPSC 78

RESULT 13
US-09-335-325-18
; Sequence 18, Application US/09335325
; Patent No. US2002016471A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid

```

```
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K3
/ SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-335-325-18

Query Match          52.1%; Score 281; DB 10; Length 78;
Best Local Similarity 56.4%; Pred. No. 4e-23;
Matches 44; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCCQSSSMIPHHQRTPENYPNAGLTENYCRNPDGKQPW 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLAGTGTENYRGVAVTVSGHTCQHSQAQTPTHTNRTPEPFCKNLNENYCRNPDGKAPW 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 68 CYTTDPCVWEYCNLTQC 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CHTNSQVRWEYCKIPSC 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-335-325-19
; Sequence 19, Application US/09335325
; Patent No. US2002016471A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
```

```
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: K3
/ SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-335-325-19

Query Match          52.1%; Score 281; DB 10; Length 78;
Best Local Similarity 57.7%; Pred. No. 4e-23;
Matches 45; Conservative 7; Mismatches 26; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCCQSSSMIPHHQRTPENYPNAGLTENYCRNPDGKQPW 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLAGTGTENYRGVAVTVSGHTCQHSQAQTPTHTNRTPEPFCKNLNENYCRNPDGKAPW 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 68 CYTTDPCVWEYCNLTQC 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CHTNSQVRWEYCKIPSC 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
US-10-131-241-18
; Sequence 18, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Fortier, Anne H.
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 18
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-18

Query Match          52.1%; Score 281; DB 15; Length 78;
Best Local Similarity 56.4%; Pred. No. 4e-23;
Matches 44; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCCQSSSMIPHHQRTPENYPNAGLTENYCRNPDGKQPW 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLAGTGTENYRGVAVTVSGHTCQHSQAQTPTHTNRTPEPFCKNLNENYCRNPDGKAPW 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 68 CYTTDPCVWEYCNLTQC 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CHTNSQVRWEYCKIPSC 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: September 29, 2003, 08:32:03
Job time : 64 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:12:27 ; Search time 40 Seconds
(without alignments)
218.784 Million cell updates/sec

Title: US-10-088-548-4
Perfect score: 539
Sequence: 1 KSPVQDCYHGDRSYRGIS.....DPCVRWEYCNLTQCSFESG 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168692 residues
Total number of hits satisfying chosen parameters: 32085

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	329	61.0	89	A60140	plasmin (EC 3.4.21
2	66	12.2	86	S20651	Ig heavy chain V r
3	57.5	10.7	77	T14395	proteinase inhibit
4	55.5	10.3	86	A37910	muscarinic toxin 2
5	55.5	10.3	88	S16161	BT14 protein - ba
6	55	10.2	86	S26459	Ig kappa chain V r
7	55	10.2	86	S29587	Ig kappa chain V r
8	54	10.0	57	T2N7BE	short toxin CM-1b
9	52.5	9.7	67	S24216	Ig kappa chain - m
10	51.5	9.6	55	D96636	unknown protein, 8
11	51	9.5	88	G69136	ribosomal protein
12	50	9.3	72	D85557	hypothetical prote
13	50	9.3	76	1HER2	high potential iro
14	49.5	9.2	82	JC5892	cobrotoxin b - Chi
15	49.5	9.2	84	PSG142	replication-associ
16	49	9.1	72	1QECFS	hypothetical 8k pr
17	49	9.1	81	1IPDK	insulin precursor
18	48.5	9.0	72	1N2EIV	long neurotoxin 1
19	48.5	9.0	77	S30578	proteinase inhibitor
20	48.5	9.0	84	E86209	Ig kappa chain V r
21	48.5	9.0	86	S78488	class II histocomp
22	48.5	9.0	89	PI0129	MHC class II histo
23	48.5	9.0	89	1G8555	kappa-3 neurotoxin
24	48	8.9	87	B60549	MHC class II histo
25	48	8.9	89	S38704	sox protein - less
26	47.5	8.8	55	S34330	short neurotoxin 3
27	47.5	8.8	62	1N1N3M	proteinase inhibit
28	47.5	8.8	77	B84433	Ig kappa chain V r
29	47.5	8.8	86	S34086	

30	47.5	8.8	87	2	S34084	Ig kappa chain V r
31	47.5	8.8	87	2	S34083	Ig kappa chain V r
32	47.5	8.8	88	2	S21528	Ig kappa chain V r
33	47.5	8.8	88	2	S21522	Ig kappa chain V r
34	47	8.7	55	2	E58892	H+-transporting tw
35	47	8.7	62	2	G25866	short neurotoxin b
36	47	8.7	66	2	P00342	spike glycoprotein
37	47	8.7	66	2	P00340	spike glycoprotein
38	47	8.7	70	2	T32946	hypothetical prote
39	47	8.7	73	2	AD1043	transcription regu
40	47	8.7	87	2	S16842	Ig kappa chain V r
41	47	8.7	87	2	S16841	Ig kappa chain V r
42	46.5	8.6	77	2	C84433	proteinase inhibit
43	46.5	8.6	81	2	A11441	intestinal trefoil
44	46	8.5	61	1	N1R12	short neurotoxin 2
45	46	8.5	76	2	T34687	hypothetical prote

ALIGNMENTS

RESULT 1

A60140
plasmin (EC 3.4.21.7) precursor - chicken (fragment)
N;Alternate names: plasminogen
C;Species: Gallus gallus (chicken)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C;Accession: A60140
R;Gyenes, M.; Patthy, L.
Biochim. Biophys. Acta 832, 326-330, 1985
A;Title: The kringle 4 domain of chicken plasminogen.
A;Reference number: A60140; WUID:8607796; PMID:4074753
A;Accession: A60140
A;Molecule type: protein
A;Residues: 1-89 <GR>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homo
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinas.
F;6-83/Domain: kringle homology <KRG>
F;6-83/27-66,55-78/Disulfide bonds: #status predicted
F;6/Binding site: carboxydrate (Asn) (covalent) #status experimental

Query Match 61.0%; Score 329; DB 2; Length 89;
Best Local Similarity 61.0%; Pred. No. 2e-25;
Matches 50; Conservative 16; Mismatches 16; Indels 0; Gaps 0;

QY	6	QDCYHGDRSYRGISSTVTGTCQSSSMIPHWORTPENTPNAGLTENTCRNPDSGKQ	65
DB	4	EECYQNGVSYRGTSFTITGKCAWNSMSPHRHNKTESHPFNADLRQNYCRNPDA	63
QY	66	PWCYTDPQVWEYCNLTQCS	87
DB	64	PWCITDPQVWEYCNLTQCS	85

RESULT 2

S20651
Ig heavy chain V region - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C;Accession: S20651; S20647
R;Losman, M.; Fasy, T.M.; Novick, K.E.; Monestier, M.
Submitted to the EMBL Data Library, February 1992
A;Description: Relationships among antinuclear antibodies from autoimmune MRL mice r
A;Reference number: S20639
A;Accession: S20651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-86 <LOS>
A;Cross-references: EMBL:X65007; NID:G52647; PIDN:CAAA6140.1; PID:G52648; EMBL:X6501
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 12.2%; Score 66; DB 2; Length 86;

Db | | | | | | : : | | | | | | | | | |

26 GSGTSY-----SLTISRMEADAATYCCQWSSYPHPVRCWDQAGAE 67

RESULT 10

D96636
unknown protein, 87272-87105 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: D96636
R:rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lib, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzi, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96636
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-55 <STO>
A:Cross-references: GB:AE005173; NID:g6751695; PIDN:AAF27678.1; GSPDB:GN00141
C:Genetics:
A:Gene: WTP1.20
A:Map position: 1
C:Superfamily: gamma-thionin

Query Match 9.6%; Score 51.5; DB 2; Length 55;
Best Local Similarity 26.6%; Pred No. 1.8e+02;
Matches 17; Conservative 7; Mismatches 15; Indels 25; Gaps 3;

QY 23 TVTGTCQSSMIPHWHORTENYP-----NAGLTENYCNPDSGKQPWCYT 70
|| |||||: |: | | | | | | | : | | : |
Db 5 TVEARTCTSNLF-----NGPCLSSNCANCVHNEGSFSDGDCR----GFRRCLC 51

QY 71 TDPC 74
| | |
Db 52 TRPC 55

RESULT 11

G69136
ribosomal protein L34 - Methanobacterium thermoautotrophicum (strain Delta H)
N:Alternate names: eukaryotic ribosomal protein L34; prokaryotic ribosomal protein L34; Methanobacterium thermoautotrophicum
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 13-Aug-1999
C:Accession: G69136
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, K. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997
A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: f
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: G69136
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-88 <MTH>
A:Cross-references: GB:AE000796; GB:AE000666; NID:g2621057; PIDN:AAB84538.1; PID:g2621057
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH29
C:Superfamily: rat ribosomal protein L34

Query Match 9.5%; Score 51; DB 2; Length 88;
Best Local Similarity 30.0%; Pred No. 3.1e+02;
Matches 15; Conservative 6; Mismatches 17; Indels 12; Gaps 2;

QY 14 RSYRGISVTVGTCQSSMIPHWHORTENYPNAGLTENYCNPDSG 63
|||: | | | | | | : : | | : | | : | | : | |

Search completed: September 29, 2003, 08:22:46
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:04:12 ; Search time 23 Seconds
(without alignments)
186.062 Million cell updates/sec

Title: US-10-088-548-4
Perfect score: 539
Sequence: 1 KSPVVQCYHGDSYRGIS.....DPCVRWEYCNLTQCSEYEG 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 11738

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Lasting first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.5	10.3	86	1 TXM2_DENAN	P18328 dendroaspis
2	55.5	10.3	88	1 CR14_HORVU	P26154 hordeum vul
3	54	10.0	57	1 TXWB_HEMHA	P01402 hemachatus
4	54	10.0	79	1 NXS6_PSETE	O9w717 pseudonaja
5	51	9.5	88	1 RL34_METH	O26137 methanobact
6	50	9.3	76	1 HP12_ECTHA	P04169 ectothiorho
7	49.5	9.2	82	1 CBTB_NAJAT	P80958 naja atra (
8	49	9.1	72	1 YED2_ECOLI	P18393 escherichia
9	49	9.1	81	1 INS_ANAPL	P01333 anas platyr
10	48.5	9.0	72	1 NXLI_DENVI	P01394 dendroaspis
11	48.5	9.0	77	1 THG1_ARATH	O39182 arabidopsis
12	48	8.9	87	1 NXL3_BUNMU	P15817 bungarus mu
13	47.5	8.8	62	1 NXS3_NAJMO	P01432 naja mossam
14	47.5	8.8	77	1 THG2_ARATH	O41914 arabidopsis
15	47	8.7	55	1 ATP8_LATCH	O03168 latimeria c
16	47	8.7	62	1 NXS8_LATLA	P10459 laticauda l
17	46.5	8.6	60	1 TX48_DENJA	P25683 dendroaspis
18	46.5	8.6	77	1 THG3_ARATH	O92ul7 arabidopsis
19	46.5	8.6	80	1 TX3A_PHRNI	P81793 phoneutria
20	46	8.5	49	1 CBL2_VISAL	P81859 viscum albu
21	46	8.5	61	1 NXS2_HEMHA	P01433 hemachatus
22	46	8.5	63	1 CXH_FEMHA	P24778 hemachatus
23	45.5	8.4	73	1 NXLI_DENVI	P01395 dendroaspis
24	45	8.3	61	1 NXS1_NAJHA	P01429 naja haje a
25	45	8.3	64	1 NXS1_BUNFA	P10808 bungarus fa
26	45	8.3	70	1 R37A_SULSO	Q972q3 sulfolobus
27	45	8.3	79	1 NXS1_PSETE	O9w7k2 pseudonaja
28	44.5	8.3	61	1 CBT4_NAJAT	O9yhw0 naja atra (
29	44	8.2	51	1 INS_ANGAN	P07454 anser anser
30	44	8.2	61	1 NXS2_NAJHH	P25675 naja haje h
31	44	8.2	68	1 MT3_BOVIN	P37359 bos taurus
32	44	8.2	76	1 NIKM_BOVIN	Q02376 bos taurus
33	44	8.2	78	1 RL28_CORGL	Q8ns15 corynebacte

34 44 8.2 79 1 NXS3_PSETE
35 44 8.2 79 1 FLIL_PLETR
36 44 8.2 82 1 CXO6_CONTE
37 43.5 8.1 72 1 NOD1_RHILIT
38 43.5 8.1 74 1 EDDF_HUMAN
39 43.5 8.1 87 1 PAGD_SALTY
40 43 8.0 61 1 NXS3_NAJHA
41 43 8.0 61 1 NXS4_NAJHA
42 43 8.0 62 1 SCX2_TITBA
43 43 8.0 62 1 SCX2_TITSE
44 43 8.0 71 1 NXL2_NAJME
45 43 8.0 76 1 TXP4_APTSC

ALIGNMENTS

RESULT 1
TXM2_DENAN
ID TXM2_DENAN STANDARD; PRT; 86 AA.
AC P18328;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Muscarinic toxin 2 precursor.
OS Dendroaspis angusticeps (Eastern green mamba).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Elapinae; Dendroaspis.
OX NCBI_TaxID=8618;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=91320365; PubMed=1862524;
RA Ducancel F., Rowan E.G., Cassar E., Harvey A.L., Menez A.,
RA Boullain J.-C.;
RT "Amino acid sequence of a muscarinic toxin deduced from the cDNA
nucleotide sequence."
RL Toxicol 29:516-520(1991).
RN [2]
RP SEQUENCE OF 22-86.
RC TISSUE=Venom;
RX MEDLINE=91320366; PubMed=1862525;
RA Karlsson E., Risinger C., Jolkonnen M., Wernstedt C., Adem A.;
RT "Amino acid sequence of a snake venom toxin that binds to the
muscarinic acetylcholine receptor."
RL Toxicol 29:521-526(1991).
CC -!- FUNCTION: BINDS TO THE MUSCARINIC ACETYLCHOLINE RECEPTOR.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: Belongs to the snake toxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; X52292; CAA36541.1; -
CC PIR; A37910; A37910.
CC HSP; P01382; INTN.
CC InterPro: IPR003572; Cytotoxin.
CC InterPro: IPR003571; Snake_toxin.
CC Pfam; PF00087; toxin; 1.
CC PRINTS; P00282; CYTOTOXIN.
CC PRODOM; PD000206; Snake_toxin; 1.
CC PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Toxin; Neurotoxin; Signal.
FT SIGNAL 1 21
FT CHAIN 22 86 MUSCARINIC TOXIN 2.
FT DISULFID 24 45 BY SIMILARITY.

```

FT DISULFID 38 63 BY SIMILARITY.
FT DISULFID 67 78 BY SIMILARITY.
FT DISULFID 79 84 BY SIMILARITY.
SQ SEQUENCE 86 AA; 9375 MW; 6F062C970074D653 CRC64;

Query Match 10.3%; Score 55.5; DB 1; Length 86;
Best Local Similarity 28.8%; Pred. No. 27;
Matches 15; Conservative 5; Mismatches 31; Indels 1; Gaps 1;

QY 24 VTGRTCSWSM-IPHWORTPENYPNAGLTENYCRNPDSGKQPCWYTTDPC 74
|| | : : || ||| | | : | | |
DB 33 VTTEDCPAGNVCFKRWVTPKNDVIKGAATCFKVDNDPIRCCGTDKC 84

RESULT 2
CR14_HORVU
ID CR14_HORVU STANDARD; PRT; 88 AA.
AC P26154;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cold-regulated protein BLT14.
GN BLT14.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Igri; TISSUE=Shoot meristem;
RA Dunn M.A., Hughes M.A., Pearce R.S., Jack P.L.;
RT "Molecular characterization of a barley gene induced by cold
treatment.";
RL J. Exp. Bot. 41:1405-1413(1990).
CC -!- INDUCTION: By cold stress.
CC -!- MISCELLANEOUS: CONTAINS SEVERAL ARGinine RESIDUES IN CLOSE
PROXIMITY WHICH MAY BE INVOLVED IN PROTEIN RNA INTERACTIONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J57554; CAA40779.1; --
DR PIR; S16161; S16161.
KW RNA-binding.
FT DOMAIN 32 41 ARG-RICH.
SQ SEQUENCE 88 AA; 9796 MW; C125D831E9BB40D0 CRC64;

Query Match 10.3%; Score 55.5; DB 1; Length 88;
Best Local Similarity 29.4%; Pred. No. 28;
Matches 20; Conservative 4; Mismatches 25; Indels 19; Gaps 4;

QY 13 GRSYVGSSTVTGRTCSWSMIPHWORTPENYPNAGLTENYCRNP-----DSGKQPCW 68
|| | | | | | | : | : | | | | |
DB 22 GRERGGCSDFRC---RCQRW-----RRRLQFGFLAAGGNRYRNKHYPAGGDPW- 69

QY 69 YTTDFCYR 76
||| |
DB 70 ---DPCYR 74

RESULT 3
TXWB_HEMHA
ID TXWB_HEMHA STANDARD; PRT; 57 AA.
AC P01402;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Weak toxin CM-1b.
OS Hemachatus haemachatus (Ringhals) (Sapedon haemachatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Hemachatus.
OX NCBI_TaxID=8626;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=80236878; PubMed=7394807;
RA Joubert F.J., Taljaard N.;
RT "The complete primary structure of toxin CM-1b from Hemachatus
haemachatus (Ringhals) snake venom.";
RL Toxicon 18:191-198(1980).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- MISCELLANEOUS: LD(50) IS 11.7 MG/KG BY INTRAVENOUS INJECTION.
CC -!- SIMILARITY: Belongs to the snake toxin family.
DR PIR; A01673; T2NJBE.
DR HSP; P01382; LNTN.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR ProDom; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; FALSE_NEG.
KW Toxin.
FT DISULFID 3 19 BY SIMILARITY.
FT DISULFID 12 37 BY SIMILARITY.
FT DISULFID 40 49 BY SIMILARITY.
FT DISULFID 50 55 BY SIMILARITY.
SQ SEQUENCE 57 AA; 6630 MW; 63DIC7818A8B2E20 CRC64;

Query Match 10.0%; Score 54; DB 1; Length 57;
Best Local Similarity 32.7%; Pred. No. 25;
Matches 17; Conservative 7; Mismatches 16; Indels 12; Gaps 4;

QY 28 TCQSWSMIPHWORTPENYPN-----AGLTENYCRNPDSGKQPCWYTTDPC 74
||| : | : || : | : || : | : || |
DB 11 TCQPEQKFC---YSDTMTFFNEFVLSGCT--FCTDSEGER--CCTTDRCC 55

RESULT 4
NXS6_PSETE
ID NXS6_PSETE STANDARD; PRT; 79 AA.
AC Q9WJ77;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Short neurotoxin 6 precursor (Alpha neurotoxin 6).
GN SNTX6.
OS Pseudonaja textilis (Eastern brown snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Acanthophiinae; Pseudonaja.
OX NCBI_TaxID=8673;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=99449602; PubMed=10518793;
RA Gong N.L., Armugam A., Jeyaseelan K.;
RT "Postsynaptic short-chain neurotoxins from Pseudonaja textilis: cDNA
cloning, expression and protein characterization.";
RL Eur. J. Biochem. 265:982-989(1999).
RN [2]
RP SEQUENCE FROM N.A. AND IDENTIFICATION OF INTRONS.
RC TISSUE=Venom gland;
RX MEDLINE=20279909; PubMed=10818230;
RA Gong N.L., Armugam A., Jeyaseelan K.;
RT "Molecular cloning, characterization and evolution of the genes
encoding a new group of short-chain alpha-neurotoxins in an Australian
elapid, Pseudonaja textilis.";
RL FEBS Lett. 473:303-310(2000).
CC -!- FUNCTION: Lethal neurotoxin, binds and inhibits nicotinic
acetylcholine receptors (nAChR) (By similarity).

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; AE000796; AAB84538.1; -.
DR PIR; G69136; G69136.
DR HAMAP; MF 00349; -. 1.
DR InterPro; IPR001284; Ribosomal_L34E.
DR Pfam; PF01199; Ribosomal_L34e; 1.
DR PRINTS; PR01250; RIBOSOMAL_L34.
DR ProDom; PD005148; Ribosomal_L34E_C; 1.
DR PROSITE; PS01145; RIBOSOMAL_L34E; 1.
KW Ribosomal protein; Complete proteome.
KW RIBOSOMAL protein; Complete proteome.
SQ SEQUENCE 88 AA; 10301 MW; 3F6C1AC5D2E41ECE CRC64;

Query Match 9.5%; Score 51; DB 1; Length 88;
Best Local Similarity 30.0%; Pred. No. 82;
Matches 15; Conservative 6; Mismatches 17; Indels 12; Gaps 2;

QY 14 RSYRGISSTVTGRFCQSWSMIPHHHQRTPEYNPAGLTENYCRNPDGS 63
||||:| | ||| : : : : | | | | |
DB 9 RSYRKFKTKPGGRT-----VTHYRRKKPSKHVCAG-----CGKPLHG 46

RESULT 6
HP12_ECTHA
ID HP12_ECTHA STANDARD; PRT; 76 AA.
AC P04169;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE High potential iron-sulfur protein isozyme II (HiPIP 2) (High-redox-
DE potential ferredoxin 2).
GN HIP2.
GN Ecotrichothodospira halophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC Ecotrichothodospiraceae; Halorhodospira.
OX NCBI_taxID=1053;
RN [1]
RP SEQUENCE.
RX MEDLINE=85305760; PubMed=4037807;
RA Tedro S.M., Meyer T.B., Kamen M.D.;
RT "Amino acid sequence of high-redox-potential ferredoxin (HiPIP)
RT isozymes from the extremely halophilic purple phototrophic bacterium,
RT Ecotrichothodospira halophila.";
RL Arch. Biochem. Biophys. 241:656-664(1985).
CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S
CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST
CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST
CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.
CC THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +50 MV.
CC -!- SUBUNIT: Homodimer (probable).
CC -!- METALENEOUS: IN E. HALOPHILIA, TWO HIPIP ISOZYMES ARE FOUND; THEY
CC HAVE THE LOWEST REDOX POTENTIALS YET DETERMINED.
CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN
CC (HIPIP) FAMILY.
DR PIR; A00271; IHER2.
DR HSP; P04168; 2HIP.
DR InterPro; IPR000170; Hipot_ironsulf.
DR Pfam; PR01355; HIPIP; 1.
DR PRINTS; PR003374; HIPIPFERDOXIN.
DR PROSITE; PS00596; HIPIP; 1.
KW Electron transport; Iron-sulfur; 4Fe-4S.
FT METAL 38 38 IRON-SULFUR (4FE-4S).
FT METAL 41 41 IRON-SULFUR (4FE-4S).
FT METAL 54 54 IRON-SULFUR (4FE-4S).
FT METAL 70 70 IRON-SULFUR (4FE-4S).
SQ SEQUENCE 76 AA; 8560 MW; 12D1524E80BAF7AC CRC64;

Query Match 9.3%; Score 50; DB 1; Length 76;
Best Local Similarity 25.5%; Pred. No. 89;

```

Query Match      9.2%;      Score 49.5;  DB 1;  Length 82;
Best Local Similarity 30.0%;  Fred. No. 1.1e-02;
Matches 18;  Conservative 24;  Mismatches 13;  Gaps 3;

QY 20 SSSVTGRTGTCQSSMSIPHW---HQTPEPNPAGLITENYCRNP--DSGKQPKVCTTDC 74
      ||| : ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 29 SSGTPPTKTCSETNYCKKWSDBRG-----IIERGGCKPVPGLVNLNCCCTDRC 80

```



```

RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=74309059; PubMed=4855008;
RA Banks B.E.C., Miledi R., Shipolini R.A.;
RT "The primary sequences and neuromuscular effects of three neurotoxic
RN polypeptides from the venom of Dendroaspis viridis.";
RL Eur. J. Biochem. 45:457-468(1974).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- PTM: NEUROTOXIN 4.7.3 DIFFERS FROM 4.9.3 ONLY IN THAT
CC TRP-26 HAS UNDERGONE PARTIAL PHOTOOXIDATION.
CC -!- MISCELLANEOUS: LD(50) IS 0.9 MG/KG BY INTRAPERITONEAL INJECTION.
CC -!- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR PIR: A01665; N2EPIV.
DR HSP: P01386; ITXA.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; toxin; 1.
DR ProDom: PD000206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; 1.
KW Toxin; Neurotoxin; Multigene family.
FT DISULFID 3 21 BY SIMILARITY.
FT DISULFID 14 42 BY SIMILARITY.
FT DISULFID 27 31 BY SIMILARITY.
FT DISULFID 46 57 BY SIMILARITY.
FT DISULFID 58 63 BY SIMILARITY.
SQ SEQUENCE 72 AA; 7970 MW; F299DA0782163B4F CRC64;

Query Match          9.0%; Score 48.5; DB 1; Length 72;
Best Local Similarity 26.3%; Pred. No. 1.2e+02;
Matches 20; Conservative 6; Mismatches 31; Indels 19; Gaps

QY      1 KSPVV--QDCYHGDRGYRGISSTVTGTCTQSWSMPIHHGHGPENYPNAGLTENVCR 58
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
       5 KTSVFPEFCPHGENICY-----TETWCADNC-----QRKREELCAAT---CP 47

QY      59 NPDSGKQPNCYITDPC 74
DB      : | : ||| |
       48 KVKGAVGIKCSTDNC 63

RESULT 11
THGL_ARATH
ID Q39182; STANDARD; PRT; 77 AA.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gamma-thionin homolog AT2G02100 precursor.
GN AT2G02100 OR FS04.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. C24; TISSUE=Flower buds;
RA Yu D.Y., Quigley F., Maché R.;
RT "Isolation and expression of a cDNA encoding protease inhibitor
RN precursor.";
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Roundsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Kechum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.
```



```

CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z17665; CAA79029.1; -
CC EMBL: AC005936; AAC97222.1; -.
CC PIR: B84433; B84433.
CC HSSP: P20230; IGPT.
CC InterPro: IPR002118; Gamma-thionin.
CC DR InterPro: IPR003614; Knot1.
CC DR Pfam: PF00304; Gamma-thionin; 1.
CC DR ProDom: PD002594; G_Purothionin; 1.
CC DR SMART: SM00505; Knot1; 1.
CC DR PROSITE: PS00940; GAMMA_THIONIN; 1.
CC KW plant defense; Signal; Multigene family.
CC FT SIGNAL 1 30 POTENTIAL.
CC FT CHAIN 31 77 GAMMA-THIONIN HOMOLOG AT2G02120.
CC FT DISULFID 33 77 BY SIMILARITY.
CC FT DISULFID 44 64 BY SIMILARITY.
CC FT DISULFID 50 71 BY SIMILARITY.
CC FT DISULFID 54 73 BY SIMILARITY.
CC FT CONFLICT 19 19 V -> D (IN REF. 1).
CC FT CONFLICT 69 69 R -> G (IN REF. 1).
CC SQ SEQUENCE 77 AA; 8578 MW; F87BE5AF727A9A45 CRC64;

Query Match 8.8%; Score 47.5; DB 1; Length 77;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 18; Conservative 4; Mismatches 28; Indels 13; Gaps 2;

QY 18 GISTTTVTGRTQSSMIPHWORTENTPNAGLTENVCNRPD-----SGQPKVCYTT 71
   I: || ||| || || : ||| : ||| : ||| : ||| : ||| :
Db 22 GMGPTVEARTCASOS-----QRFKGKCVSDTNCNVCHNEGPGGDCRGFRRCFT 74
   QY 72 DPC 74
   Db 75 RNC 77

RESULT 15
ATP8_LATCH STANDARD; PRT; 55 AA.
ID ATP8_LATCH
AC Q03168;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A5L).
DE WTAP8 OR ATP8.
GS Latimeria chalumnae (Latimeria) (Coelacanth).
OS Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Coelacanthiformes; Coelacanthidae; Latimeria.
OC NCBI_TaxID=7897;
EN [1]
RN SEQUENCE FROM N.A.
RA Zardoya R., Meyer A.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(O) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U82228; AAC60322.1; -.
CC PIR: E58892; E58892.
CC DR InterPro: IPR001421; ATPase8 mt.

```

DR Pfam; PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 4 24 POTENTIAL.
SQ SEQUENCE 55 AA; 6569 MW; 7FE36319B8AF825B CRC64;

Query Match 8.7%; Score 47; DB 1; Length 55;
Best Local Similarity 31.7%; Pred. No. 1.3e+02;
Matches 13; Conservative 6; Mismatches 16; Indels 6; Gaps 3;

QY 31 SW---SSMIPWHQRTPEKPNAGLTENYCR-NPDSGKQPW 67
|| :|:| | :|:| :|:| :|:| :|:|
Db 16 SWLFLTLMLFSTQL--HTFPNMPSTQNMCKQEPPEFTWTPW 54

Search completed: September 29, 2003, 08:20:19
Job time : 25 secs

Wed Oct 1 15:58:35 2003

us-10-088-548-4.sep29.rag

Page 9

DD 64 CHTINSQVRWEYCKTPSCS 82

Search completed: September 29, 2003, 08:19:47
Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:11:29 ; Search time 93 Seconds
(without alignments)
252.503 Million cell updates/sec

Title: US-10-088-548-4
Perfect score: 539
Sequence: 1 KSPVQDCYHGDGRSGIS.....DPCVRWEYCNLTQCSFESG 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 118357

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_prodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	236	43.8	53	4 Q8UM12	Q8um12 homo sapien
2	207.5	38.5	90	4 Q8NG20	Q8ng20 homo sapien
3	188	34.9	60	4 Q9UKJ7	Q9ukj7 homo sapien
4	120.5	22.4	75	6 Q8BGN9	Q8bgn9 bos taurus
5	109	20.2	25	4 Q9UD88	Q9ud88 homo sapien
6	62	11.5	69	5 Q8MVA6	Q8mva6 ixodes scap
7	60.5	11.2	87	12 Q8V513	Q8v513 human respi
8	59.5	11.0	82	7 Q30599	Q30599 macaca mula
9	59.5	11.0	82	7 Q19292	Q19292 macaca mula
10	59	10.9	78	6 Q8SQ84	Q8sq84 gorilla gor
11	58.5	10.9	87	12 Q8BBA5	Q8bba5 human respi
12	58.5	10.9	87	12 Q8BBA4	Q8bba4 human respi
13	58	10.8	64	13 Q8DFE1	Q8dfe1 oncorhynch
14	57.5	10.7	77	10 Q39403	Q39403 brassica ca
15	57.5	10.7	83	12 Q8V512	Q8v512 human respi
16	57.5	10.7	83	12 Q8BBA8	Q8bba8 human respi

17	57.5	10.7	87	12 Q8V314	Q8v314 human respi
18	57	10.6	85	11 Q9JHV3	Q9jhy3 mus musculu
19	56.5	10.5	80	12 Q8V516	Q8v516 human respi
20	56.5	10.5	80	12 Q9IUD1	Q9iud1 human respi
21	56.5	10.5	83	12 Q9IUE3	Q9iue3 human respi
22	56.5	10.5	83	12 Q9IH96	Q9ih96 human respi
23	56.5	10.5	83	12 Q8V5H5	Q8v5h5 human respi
24	56.5	10.5	83	12 Q8BBA3	Q8bba3 human respi
25	56.5	10.5	87	12 Q8V517	Q8v517 human respi
26	56.5	10.3	65	13 Q9PRY3	Q9pry3 dendroaspis
27	54	10.0	72	11 Q99LN5	Q99ln5 mus musculu
28	53.5	9.9	80	12 Q9IH93	Q9ih93 human respi
29	53.5	9.9	81	7 Q30728	Q30728 macaca neme
30	53.5	9.9	83	12 Q9IH91	Q9ih91 human respi
31	53.5	9.9	83	12 Q9IH88	Q9ih88 human respi
32	53.5	9.9	83	12 Q9IH92	Q9ih92 human respi
33	53.5	9.9	83	12 Q9IH87	Q9ih87 human respi
34	53.5	9.9	83	12 Q9IUD2	Q9iud2 human respi
35	53.5	9.9	83	12 Q9IMX7	Q9imx7 human respi
36	53.5	9.9	83	12 Q9IH95	Q9ih95 human respi
37	53.5	9.9	83	12 Q9IUC9	Q9iuc9 human respi
38	53	9.8	70	7 Q46870	Q46870 homo sapien
39	53	9.8	85	6 Q46556	Q46556 equus cabal
40	52.5	9.7	66	4 Q9H4U9	Q9h4u9 homo sapien
41	52.5	9.7	81	7 Q8WLN3	Q8wln3 papio anubi
42	52	9.6	76	4 Q16198	Q16198 homo sapien
43	51.5	9.6	55	10 Q9C947	Q9c947 arabidopsis
44	51.5	9.6	80	7 Q8XR03	Q8xrd3 salmo salar
45	51.5	9.6	82	7 Q9MXN5	Q9mxn5 macaca mula

ALIGNMENTS

RESULT 1

ID	Q9UMI2	PRELIMINARY;	PRT;	53 AA.
AC	Q9UMI2;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	PLG protein (fragment).			
GN	PLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85023311; PubMed=6148961;			
RA	Mallinowski D.P., Sadler J.E., Davie B.W.;			
RT	"Characterization of a complementary deoxyribonucleic acid coding for			
RL	human and bovine plasminogen."			
CC	Biochemistry 23:4243-4250(1984).			
CC	-I- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.			
DR	EMBL; K02921; AAA60123.1; -			
DR	HSSP; P00747; 2PK4			
DR	InterPro; IPR000001; Kringle.			
DR	Ffam; PF00051; Kringle; 1.			
DR	PRINTS; PR00018; KRINGLE.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00130; KR; 1.			
DR	PROSITE; PS00021; KRINGLE_1; FALSE_NEG.			
DR	PROSITE; PS50070; KRINGLE_2; 1.			
KW	Glycoprotein; Kringle.			
FT	NON_TER 1			
SQ	SEQUENCE 53 AA; 5893 MW; 2EB36E715CD40A47 CRC64;			

Query Match 43.8%; Score 236; DB 4; Length 53;
Best Local Similarity 81.8%; Pred. No. 3.4e-19;
Matches 40; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 SPVQDCYHGDGRSGISSTVTGTGTCQSSMIPHWHTPENTPNA 50


```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94060120; PubMed=7848387;
RA Praeflinger D., McLean J., Scanu A.M.;
RT "Amplification of human APO(a) kringle 4-37 from blood lymphocyte
  DNA.";
RL Biochim. Biophys. Acta 1225:107-109(1993).
DR HSP: P00747; 1KEN
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; Kringle.1.
KW Glycoprotein; Kringle.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2928 MW; BEED4C62FA480A8 CRC64;

Query Match 20.2%; Score 109; DB 4; Length 25;
Best Local Similarity 72.7%; Pred. No. 2.3e-05;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 68 CYTDPVWVWYCNLTQCSSE 89
Db 1 CFTMDPSIRWYCNLTQCSDE 22

RESULT 6
Q8MVA6 PRELIMINARY; PRT; 69 AA.
AC Q8MVA6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Putative 5.3 kDa secreted protein.
OS Ixodes scapularis (Black-legged tick) (Deer tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX NCBI_TaxID=6945;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rhode Island; TISSUE=Salivary gland;
RA Valenzuela J.G., Francischetti I.N., Pham V.M., Garfield M.,
RA Mather T.N., Ribeiro J.M.C.;
RT "Exploring the Sialome of the Tick Vector of Lyme Disease, Ixodes
  scapularis.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF483734; AAK93656.1; -
SQ SEQUENCE 69 AA; 7817 MW; EBB93FDBFA754E24 CRC64;

Query Match 11.5%; Score 62; DB 5; Length 69;
Best Local Similarity 35.0%; Pred. No. 13;
Matches 14; Conservative 5; Mismatches 11; Indels 10; Gaps 3;

QY 60 PDSGKQW-----CYTDPVWVWYCNLTQCSSE 91
Db 24 PDGP-QPWQVYKAGRPYCI-PCRKHDECRVGCSCNNG 61

RESULT 7
Q8V5I3 PRELIMINARY; PRT; 87 AA.
AC Q8V5I3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Attachment glycoprotein G (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-SA98VI53;
RX MEDLINE=21405813; PubMed=11514720;
RA "Genter M., Madhi S.A., Tiemessen C.T., Schoub B.D.;
  Genetic diversity and molecular epidemiology of respiratory syncytial
  virus over four consecutive seasons in South Africa: identification of
  new subgroup A and B genotypes.";
RL J. Gen. Virol. 82:2117-2124(2001).
DR EMBL: AF348815; AAL60316.1; -
DR InterPro: IPR000925; Glycoprot_G.
DR Pfam: PF00802; Glycoprotein_G; 1.
FT NON_TER 1 1
FT NON_TER 87 AA; 9431 MW; CEF8B4DD27344769 CRC64;
SQ SEQUENCE 87 AA; 9431 MW; CEF8B4DD27344769 CRC64;

Query Match 11.2%; Score 60.5; DB 12; Length 87;
Best Local Similarity 32.1%; Pred. No. 24;
Matches 17; Conservative 4; Mismatches 27; Indels 5; Gaps 1;

QY 1 KSPVQDCYHGDRSYRGISSTVTGRTCCQSSSMIPHWQRTPENYPNAGLT 53
Db 21 KKPTPKTERTGSTSQSTVLDTTTSKHTIQQS-----LHSITPTNPSTQT 68

RESULT 8
Q30599 PRELIMINARY; PRT; 82 AA.
AC Q30599;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Major histocompatibility complex class II (Fragment).
GN MHC-DRB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97152405; PubMed=8995183;
RA Knapp J.A., Cadavid L.F., Eberle M.E., Knechtle S.J., Bontrop R.E.,
RA Watkins D.I.;
RT "Identification of new manu-DRB alleles using DGGE and direct
  sequencing.";
RL Immunogenetics 45:171-179(1997).
DR EMBL: U57941; AAC50974.1; -
DR InterPro: IPR000353; MHC_II_beta.
DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1 1
FT NON_TER 82 82
SQ SEQUENCE 82 AA; 9945 MW; 1FE11A6520A60108 CRC64;

Query Match 11.0%; Score 59.5; DB 7; Length 82;
Best Local Similarity 30.8%; Pred. No. 29;
Matches 16; Conservative 8; Mismatches 17; Indels 11; Gaps 3;

QY 9 YHGDRSYRGISSTVTGRTCCQSSSMIPHWQRTPENYPNAGLTENYCRN 59
Db 33 FDSVDVGEYRAVSE---LGRPTAESWNSQKDYLEQ-----RGQVDNYCRH 74

RESULT 9
Q19292 PRELIMINARY; PRT; 82 AA.
AC Q19292;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II DR beta (fragment).
GN HLA-DRB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE OF 3-30 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=94179810; PubMed=8133043;
RA Sliedrecht B.L., Oetting N., van Besouw N., Jonker M., Bontrop R.E.;
RT "Expansion and contraction of rhesus macaque DRB regions by
RT duplication and deletion.";
RL J. Immunol. 152:2298-2307(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA de Groot N.N.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031268; AAB87440.1; -.
DR HSSP; P13758; IDLH;
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9960 MW; 71FA3A6520A60119 CRC64;

Query Match 11.0%; Score 59.5; DB 7; Length 82;
Best Local Similarity 30.8%; Pred. No. 29;
Matches 16; Conservative 8; Mismatches 17; Indels 11; Gaps 3;

QY 9 YHGDGRSYRGISSTVTGGR-TCQSWSSMIPHWHQRTPEPNAGLTENYCRN 59
DB 33 FDSVDGEYRAVSE--LGRPTAESWNSQDYLEQKRAE-----VDNYCRH 74

RESULT 10
Q8SQ84 ID Q8SQ84 PRELIMINARY; PRT; 78 AA.
AC Q8SQ84;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE MHC-G (Fragment).
GN MHC-G.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97364681; PubMed=9218527;
RA Arnaiz-Villena A., Martinez-Laso J., Alvarez M., Castro M.J.,
RA Varela P., Gomez-Casado E., Suarez B., Recio M.J., Vargas-Alarcon G.,
RA Morales P.;
RT "Primate Mhc-B and -G alleles.";
RL Immunogenetics 46:251-266(1997).
DR EMBL; I41256; AAL77572.1; -.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom; PD000050; MHC_I; 1.
FT NON_TER 1
FT NON_TER 78
SQ SEQUENCE 78 AA; 8891 MW; EC12D3F3A58FC7B3 CRC64;

Query Match 10.9%; Score 59; DB 6; Length 78;
Best Local Similarity 22.7%; Pred. No. 31;
Matches 15; Conservative 10; Mismatches 37; Indels 4; Gaps 1;

QY 12 DGRSYRGISSTVTGRTQCSWSSMIPHWHQRTPEPNAGLTENYCRNPDSGKQPCWYTT 71
DB 7 DGRLLGEYQYADGKDYLALNEDLRSWTAADTA----AQISKKKCEAPNAPQRRAYLE 62

QY 72 DPCVRW 77
DB 63 GTCVEW 68

RESULT 11
Q8BBA5 ID Q8BBA5 PRELIMINARY; PRT; 87 AA.
AC Q8BBA5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AGK6-99;
RX MEDLINE=22213634; PubMed=12226836;
RA Venter M., Collinson M., Schoub B.D.;
RT "Molecular epidemiological analysis of community circulating
RT respiratory syncytial virus in rural South Africa: Comparison of
RT viruses and genotypes responsible for different disease
RT manifestations.";
RL J. Med. Virol. 68:452-461(2002).
DR EMBL; AF548806; AAN62463.1; -.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9431 MW; F72B9DC59C975EE0 CRC64;

Query Match 10.9%; Score 58.5; DB 12; Length 87;
Best Local Similarity 32.7%; Pred. No. 40;
Matches 16; Conservative 3; Mismatches 25; Indels 5; Gaps 1;

QY 1 KSPVQDCYHGDGRSYRGISSTVTGRTQCSWSSMIPHWHQRTPEPNYPN 49
DB 21 KKPTKTERGTSTQSTVLDTTTSHKTIQQQS-----LHSTTPENTPN 64

RESULT 12
Q8BBA4 ID Q8BBA4 PRELIMINARY; PRT; 87 AA.
AC Q8BBA4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AGJ16.99;
RX MEDLINE=22213634; PubMed=12226836;
RA Venter M., Collinson M., Schoub B.D.;
RT "Molecular epidemiological analysis of community circulating
RT respiratory syncytial virus in rural South Africa: Comparison of
RT viruses and genotypes responsible for different disease
RT manifestations.";
RL J. Med. Virol. 68:452-461(2002).
DR EMBL; AF548807; AAN62464.1; -.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9491 MW; 3BFB3AD57345C68 CRC64;

Query Match 10.9%; Score 58.5; DB 12; Length 87;
Best Local Similarity 32.7%; Pred. No. 40;
Matches 16; Conservative 3; Mismatches 25; Indels 5; Gaps 1;

QY 1 KSPVQDCYHGDGRSYRGISSTVTGRTQCSWSSMIPHWHQRTPEPNYPN 49

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:23:53 ; Search time 82 Seconds
(without alignments)
172.276 Million cell updates/sec

Title: US-10-088-548-6

Perfect score: 523

Sequence: 1 VRQYHGNGSQYRGFTFTV.....SIRWEYCNLTRCSDTEGTVV 89

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 693114

Minimum DB seq length: 0

Maximum DB seq length: 89

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523	100.0	89	22	AA1980
2	430	82.2	84	21	AA1981
3	427	81.6	82	10	AA1982
4	421	80.5	83	21	AA1983
5	417	79.7	78	17	AA1984
6	369	70.6	78	17	AA1985
7	349.5	66.8	86	23	AA1986
8	313	59.8	78	17	AA1987
9	299	57.2	78	17	AA1988

10	286	54.7	83	21	AA1989
11	285	54.5	78	17	AA1990
12	283	54.1	78	17	AA1991
13	281	53.7	78	17	AA1992
14	277	53.0	78	17	AA1993
15	276	52.8	83	21	AA1994
16	274	52.4	78	17	AA1995
17	271.5	51.9	89	20	AA1996
18	271.5	51.9	89	20	AA1997
19	271.5	51.9	89	20	AA1998
20	269.5	51.5	82	21	AA1999
21	269.5	51.5	84	21	AA2000
22	269.5	51.5	84	21	AA2001
23	268.5	51.3	79	17	AA2002
24	267.5	51.1	79	17	AA2003
25	264	50.5	78	17	AA2004
26	264	50.5	85	21	AA2005
27	262.5	50.2	79	9	AA2006
28	262.5	50.2	79	17	AA2007
29	262	50.1	78	17	AA2008
30	261	49.9	78	21	AA2009
31	260	49.7	78	17	AA2010
32	256.5	49.0	79	17	AA2011
33	256	48.9	81	10	AA2012
34	254.5	48.7	78	10	AA2013
35	253.5	48.5	84	21	AA2014
36	247.5	47.3	79	17	AA2015
37	245.5	46.9	79	17	AA2016
38	244	46.7	79	18	AA2017
39	240	45.9	85	21	AA2018
40	236	45.1	82	21	AA2019
41	224	42.8	86	22	AA2020
42	223	42.6	52	22	AA2021
43	223	42.6	52	22	AA2022
44	223	42.6	52	22	AA2023
45	223	42.6	52	22	AA2024

ALIGNMENTS

RESULT 1

AA1980

ID AA1980 standard; Protein; 89 AA.

AC AA1980

XX AA1980

DT 13-JUN-2001 (first entry)

XX Human apolipoprotein(a) kringle domain IV37, LK7 protein.

Human; angiogenesis inhibitor; LK7; apolipoprotein(a) kringle domain; angiotensin-mediated disease; cancer; rheumatoid arthritis; therapy; cytostatic; antirheumatic; antiarthritic; antipsoriatic; psoriasis; ocular angiogenic disease; endothelial cell proliferation; tumour; cell migration.

OS Homo sapiens.

XX WO200119868-A1.

PN 22-MAR-2001.

XX 15-SEP-1999; 99WO-KR00554.

XX 15-SEP-1999; 99WO-KR00554.

XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.

XX Chang J, Kim JS, Park EJ, Yum J, Chung S;

PI WPI; 2001-244787/25.

DR N-PSDB; AAD03257.

```

XX Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
PT diseases, e.g. cancer and rheumatoid arthritis, has human
PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
PT sequence
XX
PS Claim 2; Page 45-46; 50pp; English.
XX
CC The present sequence is human LK7 protein which contains the amino
CC acid sequence of human apolipoprotein(a) kringle domain IV37
CC (KIV37). The human apolipoprotein(a) kringle domains IV36 (LK6
CC protein), IV37 (LK7 protein) and V38 (LK8 protein) together form the
CC angiogenesis inhibitor, LK68 protein. LK68, LK6, LK7 and LK8 are
CC inhibitors are of endothelial cell proliferation, cell migration and
CC normal development of capillaries in the chick embryo chorioallantoic
CC membrane (CAM). LK68 protein, its single kringle or their functional
CC equivalents, are useful for treating angiogenesis-mediated diseases,
CC such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic
CC disease in animals or humans. LK68 is useful as an anticancer agent and
CC also for inhibiting primary tumour growth.
XX
SQ Sequence 89 AA;
Query Match 100.0%; Score 523; DB 22; Length 89;
Best Local Similarity 100.0%; Pred. No. 5.2e-46;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRQCYHGNGQSYRGTFSTVTGTCQSSMTPHRHQTPTENPNDGLTMNCRNPDADT 60
DB 1 VRQCYHGNGQSYRGTFSTVTGTCQSSMTPHRHQTPTENPNDGLTMNCRNPDADT 60
QY 61 GPWCFTTDPDSIRWEYCNLTRCSDTGTGVV 89
DB 61 GPWCFTTDPDSIRWEYCNLTRCSDTGTGVV 89
RESULT 2
AAAY77720
ID AAAY77720 standard; protein; 84 AA.
XX
AC AAAY77720;
XX
DT 12-MAY-2000 (first entry)
XX
DE Human plasminogen kringle 4 sequence.
XX
KW Kringle 4; plasminogen; anti-angiogenic; angiogenesis; angiotensin;
KW cytostatic; antiarthritic; antirheumatic; antidiabetic; ophthalmological;
KW immunosuppressant; vasotropic; vulnery; antiarteriosclerotic; human;
KW dermatological; cancer; tumour; birth control; vascularization.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 59..75
FT /note= "specifically claimed fragment (AAAY77723)"
XX
PN W0200003726-A1.
XX
PD 27-JAN-2000.
XX
PF 07-JUL-1999; 99WO-US15271.
XX
PR 14-JUL-1998; 98US-0092831.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Ji R, Trail PA;
XX
DR WPI; 2000-171200/15.
XX
PT Novel lysine binding fragments angiotensin used as antiangiogenic
PT agents in the treatment of cancer, diabetic retinopathy, rheumatoid

```

```

PT arthritis, psoriasis, atherosclerotic plaque formation, and other
PT angiogenesis diseases
XX
PS Disclosure; Fig 3; 30pp; English.
XX
CC The invention provides fragments of kringle 1, 2 or 4 of human
CC plasminogen that contain a lysine binding site and have anti-angiogenic
CC activity. The peptides of the invention function as antiangiogenic
CC agents, for the treatment process and diseases involving angiogenesis.
CC Such diseases include cancers such as solid tumours, blood born tumours
CC such as leukemias, tumours metastasis, benign tumours, such as
CC hemangiomas, acoustic acromas, neurofibromas, trachomas and pyogenic
CC granulomas, rheumatoid arthritis, ocular angiogenic diseases such as
CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
CC rubeosis, Scler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularization, telangiectasia, hemophilic joints, angiofibroma,
CC and wound granulation. The fragments are also useful in treatment of
CC disease of excessive or abnormal stimulation of endothelial cells. These
CC diseases include intestinal adhesions, atherosclerosis, scleroderma, and
CC hypertrophic scars. The fragments can be used as birth control agents by
CC preventing vascularization required for embryo transplantation. The
CC present sequence represents the kringle 4 sequence of human plasminogen
CC (angiotensin).
XX
SQ Sequence 84 AA;
Query Match 82.28; Score 430; DB 21; Length 84;
Best Local Similarity 83.38; Pred. No. 1.5e-36;
Matches 70; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 1 VRQCYHGNGQSYRGTFSTVTGTCQSSMTPHRHQTPTENPNDGLTMNCRNPDADT 60
DB 1 VQDQYHGNGQSYRGTFSTVTGTCQSSMTPHRHQTPTENPNDGLTMNCRNPDADK 60
QY 61 GPWCFTTDPDSIRWEYCNLTRCSDT 84
DB 61 GPWCFTTDPDSIRWEYCNLTRCSDT 84
RESULT 3
AAP93231
ID AAP93231 standard; peptide; 82 AA.
XX
AC AAP93231;
XX
DT 25-MAR-2003 (updated)
DT 03-APR-1990 (first entry)
XX
DE Plasminogen kringle 4 domain (residues 354-435).
XX
KW Plasminogen; activator; t-PA; fibrin; kringle domain;
KW SCU-PA;
XX
OS Homo sapiens.
XX
PN W08910401-A.
XX
PD 02-NOV-1989.
XX
PF 23-MAR-1989; 89WO-US01255.
XX
PR 22-APR-1988; 88US-0184823.
XX
PA (COLB ) COLLABORATIVE RES INC.
XX
PI Mao JI, Abercrombie DM;
XX
DR WPI; 1989-339965/46.
XX
DR N-PSDB; AAN92238.
XX
PT Modified plasminogen activator - having greater fibrin
PT selectivity and circulating half-life.

```

XX Disclosure; fig 3b; 80pp; English.

XX The sequence is encoded by a synthetic oligonucleotide and is the kringle

CC 4 domain of plasminogen. It is used to construct a modified plasminogen

CC activator (mPA) of greater fibrin selectivity than normal PA. The mPA is

CC used in the treatment of pulmonary embolism, thrombosis, myocardial

CC infarct and strokes.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 82 AA;

XX Query Match 81.6%; Score 427; DB 10; Length 82;

XX Best Local Similarity 85.2%; Pred. No. 3.1e-36;

XX Matches 69; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 VRCYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEPNYNDGLTNNYCRNPDAKT 60

DB 1 VRCYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEPNYNDGLTNNYCRNPDAKT 60

QY 2 VQCYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEPNYNDGLTNNYCRNPDAKT 61

DB 2 VQCYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEPNYNDGLTNNYCRNPDAKT 61

QY 61 GPWCFTTDPSSIRWEYCNLTRC 81

DB 62 GPWCFTTDPSSIRWEYCNLTRC 82

RESULT 4

AA08414

ID AA08414 standard; Protein: 83 AA.

XX AC AA08414;

XX DT 20-DEC-2000 (first entry)

XX DE Amino acid sequence of kringle 4 of human plasminogen.

XX KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;

XX KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;

XX KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;

XX KW adult respiratory distress syndrome; Castelman's disease; psoriasis;

XX KW hepatitis; aneurysm; renal disease; haemangioma.

XX OS Homo sapiens.

XX PN WO200049871-A1.

XX PD 31-AUG-2000.

XX PF 24-FEB-2000; 2000WO-US04798.

XX PR 24-FEB-1999; 99US-0121341.

XX PR 25-FEB-1999; 99US-0121633.

XX PR 18-NOV-1999; 99US-0166176.

XX PA (FORD-) FORD HEALTH SYSTEM HENRY.

XX PI Dou D, Chopp M, Wang L, Mikkelsen T;

XX WPI; 2000-572016/53.

XX Use of kringle protein and kringle derived from plasminogen and

PT composition comprising kringle proteins for treating tumor and

PT atherosclerosis, arthritis and retinopathy.

XX Disclosure; Fig 6; 163pp; English.

XX The specification describes a human polypeptide which is a potent

CC angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle

CC proteins, or a kringle derived from human tissue plasminogen activator

CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED

CC or tPA are useful for treating tumours, as well as atherosclerosis,

CC arthritis, retinopathy and other similar diseases. KED is also useful

CC for the treatment of diseases such as bronchial vascular congestion,

CC inflammatory bowel disease, adult respiratory distress syndrome,

CC Castelman's disease, psoriasis, hepatitis, aneurysm, renal disease

CC and haemangioma. The present sequence represents kringle 4 of human

CC plasminogen, which is used in the course of the invention.

XX Sequence 83 AA;

XX Query Match 80.5%; Score 421; DB 21; Length 83;

XX Best Local Similarity 84.1%; Pred. No. 1.3e-35;

XX Matches 69; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 VRCYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEPNYNDGLTNNYCRNPDAKT 60

DB 1 VQCYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEPNYNDGLTNNYCRNPDAKT 60

QY 61 GPWCFTTDPSSIRWEYCNLTRC 82

DB 61 GPWCFTTDPSSIRWEYCNLTRC 82

RESULT 5

AA07560

ID AA07560 standard; protein; 78 AA.

XX AC AA07560;

XX DT 22-JUN-1997 (first entry)

XX DE Human kringle 4.

XX KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;

XX KW macular degeneration; diabetic retinopathy.

XX OS Homo sapiens.

XX PN WO9635774-A2.

XX PD 14-NOV-1996.

XX PF 26-APR-1996; 96WO-US058556.

XX PR 08-MAR-1996; 96US-0612788.

XX PR 26-APR-1995; 95US-0429743.

XX PR 22-FEB-1996; 96US-0605598.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;

XX WPI; 1996-518662/51.

XX Use of angiotatin fragments or aggregates - for inhibiting

PT endothelial cell proliferation and treating angiogenesis-mediated

PT diseases, e.g. cancer, arthritis or diabetic retinopathy

XX Claim 4; Page 116; 203pp; English.

XX The invention relates to new methods and compositions for

CC inhibiting endothelial cell proliferation, using as active component

CC an angiotatin fragment, a combination of angiotatin fragments, or

CC aggregate angiotatin. The fragment is preferably derived from murine,

CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,

CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle

CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotatin has a Mol. Wt.

CC of 45-65 kD and is derived from a plasminogen fragment beginning at

CC approximately amino acid number 98 of murine, human, Rhesus, porcine or

CC bovine plasminogen. The active component can be used for treating

CC angiogenesis-mediated diseases such as cancer, arthritis, macular

CC degeneration and diabetic retinopathy. It can also be used to develop

CC antibodies for use in diagnosis, detection and therapy.

CC The present sequence, human kringle 4, is a specific angiotatin

CC fragment which can be used in the invention.

XX Sequence 78 AA;

Query Match 79.78; Score 417; DB 17; Length 78;
Best Local Similarity 85.98; Pred. No. 3e-35; 6; Indels 0; Gaps 0;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTTTGTCQSSMTPHRHQRTPEYNDGLTMYCRNPDA7GPW 63
DB 1 CYHNGQSYRGFTSTTTGTCQSSMTPHRHQRTPEYNDGLTMYCRNPDA7GPW 60

QY 64 CFTTDPISRWEYCNLRK 81
DB 61 CFTTDPISRWEYCNLRK 78

RESULT 6
AAW07559
ID AAW07559 standard; protein; 78 AA.
AC AAW07559;
XX
XX
DT 22-JUN-1997 (first entry)
DE Murine kringle 4.
XX
XX
KW angiotensin; plasminogen; kringle; angiogenesis; cancer; arthritis;
KW macular degeneration; diabetic retinopathy.
XX
XX
OS Mus musculus.
XX
XX
XX W09635774-R2.
XX
XX
PD 14-NOV-1996.
XX
XX
XX 26-APR-1996; 96WO-US05856.
XX
XX
XX 08-MAR-1996; 96US-0612788.
XX
XX 26-APR-1995; 95US-0429743.
XX
XX 22-FEB-1996; 96US-0605598.
XX
XX
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX
XX
XX Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
XX
XX WPT; 1996-518662/51.
XX
XX
XX Use of angiotensin fragments or aggregates - for inhibiting
XX endothelial cell proliferation and treating angiogenesis-mediated
XX diseases, e.g. cancer, arthritis or diabetic retinopathy
XX
XX
XX Claim 4; Page 115; 203pp; English.
XX
XX
XX The invention relates to new methods and compositions for
XX inhibiting endothelial cell proliferation, using as active component
XX an angiotensin fragment, a combination of angiotensin fragments, or
XX an aggregate angiotensin. The fragment is preferably derived from murine,
XX human, rhesus, porcine or bovine plasminogen and is a kringle 1,
XX kringle 2, kringle 3, kringle 1-3, kringle 1-2, kringle
XX 1-4 or kringle 1-4BXLs protein. The aggregate angiotensin has a Mol. wt.
XX of 45-65 kD and is derived from a plasminogen fragment beginning at
XX approximately amino acid number 98 of murine, human, rhesus, porcine or
XX bovine plasminogen. The active component can be used for treating
XX angiogenesis-mediated diseases such as cancer, arthritis, macular
XX degeneration and diabetic retinopathy. It can also be used to develop
XX antibodies for use in diagnosis, detection and therapy.
XX The present sequence, murine kringle 4, is a specific angiotensin
XX fragment which can be used in the invention, and represents amino
XX acids 377-454 of murine plasminogen.
XX
XX
XX Sequence 78 AA;

Query Match 70.68; Score 369; DB 17; Length 78;
Best Local Similarity 73.18; Pred. No. 2.4e-30;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTTTGTCQSSMTPHRHQRTPEYNDGLTMYCRNPDA7GPW 63
DB 1 CYHNGQSYRGFTSTTTGTCQSSMTPHRHQRTPEYNDGLTMYCRNPDA7GPW 60

QY 64 CFTTDPISRWEYCNLRK 81
DB 61 CFTTDPISRWEYCNLRK 78

RESULT 7
ABP02801
ID ABP02801 standard; Protein; 86 AA.
XX
XX
XX AC ABP02801;
XX
XX
DT 24-JUN-2002 (first entry)
DE
DE Human ORFX protein sequence SEQ ID NO:5584.
XX
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis.
XX
XX
XX Homo sapiens.
XX
XX
XX W0200192523-A2.
XX
XX
XX 06-DEC-2001.
XX
XX
XX 29-MAY-2001; 2001WO-US10836.
XX
XX
XX 30-MAY-2000; 2000US-206132P.
XX
XX 29-AUG-2000; 2000US-228716P.
XX
XX
XX (CURA-) CURAGEN CORP.
XX
XX
XX Shimkets RA, Leach MD;
XX
XX
XX WPI; 2002-106308/14.
XX
XX
XX N-PSDB; ABN18553.
XX
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and autoimmune disorders -
XX
XX
XX Disclosure; SEQ ID 5584; 1037pp; English.

The present invention describes substantially purified human proteins
(referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
in the specification)). ABN15762 to ABN27252 encode the human ORFX
proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
treating or preventing a pathology associated with an ORFX-associated
disorder in humans, and in the manufacture of a medicament for treating a
syndrome associated with ORFX-associated disorder. ORFX polynucleotide
sequences can be used in gene therapy. ORFX sequences can be used in the
treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
osteoarthritis, neurodegenerative disorders, disorders related to organ
transplantation, cardiovascular diseases, diabetes mellitus, systemic
lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
storage disease, various immune deficiencies and disorders, infectious
diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
disease and autoimmune inflammatory eye disease. ORFX proteins are also
useful for treating burns, incisions, ulcers, for treating osteoporosis,
bone degenerative disorders, or periodontal disease, and for gut
protection or regeneration and treatment of lung or liver fibrosis,

The invention relates to new methods and compositions for inhibiting endothelial cell proliferation, using as active component an angiotensin fragment, a combination of angiotensin fragments, or aggregate angiotensin. The fragment is preferably derived from murine, human, rhesus, porcine or bovine plasminogen and is a kringle 1, kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle 1-4 or kringle 1-4&3&5 protein. The aggregate angiotensin has a Mol. wt. of 45-65 kD and is derived from a plasminogen fragment beginning at approximately amino acid number 98 of murine, human, rhesus, porcine or bovine plasminogen. The active component can be used for treating angiogenesis-mediated diseases such as cancer, arthritis, macular degeneration and diabetic retinopathy. It can also be used to develop antibodies for use in diagnosis, detection and therapy. The present sequence, murine kringle 3, is a specific angiotensin fragment which can be used in the invention, and represents amino

30 sequences / 10 AT

CC Note: This sequence is not shown in the specification, but is deriv

QY 64 CFTTDPESIRWEYCNLTRC 81
 DB 61 CTTTSEVRWEYCTIPSC 78

RESULT 14
 AAW07553
 ID AAW07553 standard; protein; 78 AA.
 XX AC AAW07553;
 XX DE 21-JUN-1997 (first entry)
 XX DE Bovine kringle 2.
 XX KW angiotensin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 KW macular degeneration; diabetic retinopathy.
 XX OS Bos taurus.
 XX PN W09635774-A2.
 XX PD 14-NOV-1996.
 XX PF 26-APR-1996; 96WO-US05856.
 XX PR 08-MAR-1996; 96US-0612788.
 XX PR 26-APR-1996; 96US-0429743.
 XX PR 22-FEB-1996; 96US-0605598.
 XX PA (CHIL-) CHILDRENS MEDICAL CENT.
 XX PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
 XX WPT; 1996-518662/51.
 XX Use of angiotensin fragments or aggregates - for inhibiting
 PT endothelial cell proliferation and treating angiogenesis-mediated
 PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX Claim 4; Page 110; 203pp; English.
 XX The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotensin fragment, a combination of angiotensin fragments, or
 CC aggregate angiotensin. The fragment is preferably derived from murine,
 CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotensin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, bovine kringle 2, is a specific angiotensin
 CC fragment which can be used in the invention, and represents amino
 CC acids 88-165 of bovine angiotensin.
 XX Sequence 78 AA;
 Query Match 53.0%; Score 277; DB 17; Length 78;
 Best Local Similarity 52.6%; Pred. No. 6.1e-21;
 Matches 41; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTFTVIGRTCSWSMTPHRHQRTPENYNDGLTMVYCRNPDADTGP 63
 DB 1 CMFCSGENYEGKIATMSGRDCQAWDSQSPHAGTIPSKFNKLNKMYCRNPDGPRPW 60

QY 64 CFTTDPESIRWEYCNLTRC 81
 DB 61 CFTTDPQKWEFCDIPTC 78

RESULT 15
 AAB08412
 ID AAB08412 standard; Protein; 83 AA.
 XX AC AAB08412;
 XX DT 20-DEC-2000 (first entry)
 XX DE Amino acid sequence of kringle 3 of human plasminogen.
 XX KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
 KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;
 KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;
 KW adult respiratory distress syndrome; Castleman's disease; psoriasis;
 KW hepatitis; aneurysm; renal disease; haemangioma.
 XX OS Homo sapiens.
 XX PN W0200049871-A1.
 XX PD 31-AUG-2000.
 XX PF 24-FEB-2000; 2000WO-US04798.
 XX PR 24-FEB-1999; 99US-0121341.
 XX PR 25-FEB-1999; 99US-0121633.
 XX PR 18-NOV-1999; 99US-0166176.
 XX PA (FORD-) FORD HEALTH SYSTEM HENRY.
 XX PI Dou D, Chopp M, Wang L, Mikkelsen T;
 XX WPT; 2000-572016/53.
 XX Use of kringle protein and kringle derived from plasminogen and
 PT composition comprising kringle proteins for treating tumor and
 PT atherosclerosis, arthritis and retinopathy
 XX Disclosure; Fig 6; 163pp; English.
 XX The specification describes a human polypeptide which is a potent
 CC angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle
 CC proteins, or a kringle derived from human tissue plasminogen activator
 CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
 CC or tPA are useful for treating tumours, as well as atherosclerosis,
 CC arthritis, retinopathy and other similar diseases. KED is also useful
 CC for the treatment of diseases such as bronchial vascular congestion,
 CC inflammatory bowel disease, adult respiratory distress syndrome,
 CC Castleman's disease, psoriasis, hepatitis, aneurysm, renal disease
 CC and haemangioma. The present sequence represents kringle 3 of human
 CC plasminogen, which is used in the course of the invention.
 XX Sequence 83 AA;
 Query Match 52.8%; Score 276; DB 21; Length 83;
 Best Local Similarity 55.0%; Pred. No. 8.3e-21;
 Matches 44; Conservative 8; Mismatches 28; Indels 0; Gaps 0;

QY 3 QCYHNGQSYRGFTFTVIGRTCSWSMTPHRHQRTPENYNDGLTMVYCRNPDADTGP 62
 DB 3 QCLKPGGENYGRNNAVIVSGTCTQHWSAQTPTHTNRTPEFPCCKNLNDENYCRNPDGKRAP 62

QY 63 WCFTTDPESIRWEYCNLTRCS 82
 DB 63 WCHTNSQVRWEYCKTPSCS 82

Search completed: September 29, 2003, 08:33:33
 Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:32:08 : Search time 28 Seconds
(without alignments)
134.488 Million cell updates/sec

Title: US-10-088-548-6

Perfect score: 523

Sequence: 1 VRCYHGNGQSYRGTSFTTV.....SIRWEYCNLTRCSDTGTFVW 89

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 223717

Minimum DB seq length: 0
Maximum DB seq length: 89

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/ECTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	430	82.2	84	4	US-09-348-953-2
2	417	79.7	78	2	US-08-612-788-23
3	417	79.7	78	2	US-08-763-528A-5
4	417	79.7	78	3	US-09-066-028-23
5	417	79.7	78	4	US-09-335-325-23
6	369	70.6	78	2	US-08-612-788-22
7	369	70.6	78	3	US-09-066-028-22
8	369	70.6	78	4	US-09-335-325-22
9	313	59.8	78	2	US-08-612-788-17
10	313	59.8	78	3	US-09-066-028-17
11	313	59.8	78	4	US-09-335-325-17
12	299	57.2	78	2	US-08-612-788-20
13	299	57.2	78	3	US-09-066-028-20
14	299	57.2	78	4	US-09-335-325-20
15	285	54.5	78	2	US-08-612-788-18
16	285	54.5	78	3	US-09-066-028-18
17	285	54.5	78	4	US-09-335-325-18
18	283	54.1	78	2	US-08-612-788-19
19	283	54.1	78	3	US-09-066-028-19
20	283	54.1	78	4	US-09-335-325-19
21	281	53.7	78	2	US-08-612-788-21
22	281	53.7	78	3	US-08-763-528A-4
23	281	53.7	78	4	US-09-066-028-21
24	281	53.7	78	2	US-09-335-325-21
25	277	53.0	78	4	US-08-612-788-16
26	277	53.0	78	3	US-09-066-028-16
27	277	53.0	78	4	US-09-335-325-16

```

28 274 52.4 78 2 US-08-612-788-15 Sequence 15, Appl
29 274 52.4 78 3 US-09-066-028-15 Sequence 15, Appl
30 274 52.4 78 4 US-09-335-325-15 Sequence 15, Appl
31 269.5 51.5 82 4 US-09-348-953-1 Sequence 1, Appl
32 268.5 51.3 79 2 US-08-612-788-9 Sequence 9, Appl
33 268.5 51.3 79 3 US-09-066-028-9 Sequence 9, Appl
34 268.5 51.3 79 4 US-09-335-325-9 Sequence 9, Appl
35 267.5 51.1 79 2 US-08-612-788-8 Sequence 8, Appl
36 267.5 51.1 79 2 US-08-763-528A-2 Sequence 2, Appl
37 267.5 51.1 79 3 US-09-066-028-8 Sequence 8, Appl
38 267.5 51.1 79 4 US-09-335-325-8 Sequence 8, Appl
39 264 50.5 78 2 US-08-612-788-13 Sequence 13, Appl
40 264 50.5 78 3 US-09-066-028-13 Sequence 13, Appl
41 264 50.5 78 4 US-09-335-325-13 Sequence 13, Appl
42 262 50.1 78 2 US-08-612-788-14 Sequence 14, Appl
43 262 50.1 78 3 US-09-066-028-14 Sequence 14, Appl
44 262 50.1 78 4 US-09-335-325-14 Sequence 14, Appl
45 261 49.9 78 2 US-08-763-528A-3 Sequence 3, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-348-953-2
; Sequence 2, Application US/09348953
; Patent No. 6538103
; GENERAL INFORMATION:
; APPLICANT: Ji, Richard Wei-Dong
; APPLICANT: Trail, Pamela A.
; TITLE OF INVENTION: LYSINE BINDING FRAGMENTS OF ANGIOSTATIN
; FILE REFERENCE: DB11Sequences
; CURRENT APPLICATION NUMBER: US/09/348,953
; CURRENT FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/092,831
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 84
; TYPE: PRT
; ORGANISM: human
US-09-348-953-2

```

```

Query Match      82.2%; Score 430; DB 4; Length 84;
Best Local Similarity 83.3%; Pred. No. 1.1e-41;
Matches 70; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1 VRCYHGNGQSYRGTSFTTVTGRTCSWSMTPHRRHQRTPENYPNDGLTMYNCRNPDAIT 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 VQCYHGDSYRGTSSTTTTGKCCQSWSMTPHRRHQKTPENYNAGLTMYNCRNPDAK 60

```

```

QY 61 GPWCFTTDPISIRWEYCNLTRCSDT 84
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 GPWCFTTDPISIRWEYCNLTRCSGT 84

```

```

RESULT 2
US-08-612-788-23
; Sequence 23, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia

```

COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE: ORGANISM: Homo sapiens
IMMEDIATE SOURCE: CLONE: K4
US-08-612-788-23
Query Match 79.7%; Score 417; DB 2; Length 78;
Best Local Similarity 85.9%; Pred. No. 3.1e-40;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
Qy 4 CYHNGQSYRGTSFTVTGTCQSWSSMTPHRRHQRTPEYFNDGLTMNYCRNPDAQTGPW 63
Db 1 CYHGDGQSYRGTSFTVTGTCQSWSSMTPHRRHQRTPEYFNDGLTMNYCRNPDAQTGPW 60
Qy 64 CFTTDPDSIRWEYCNLRK 81
Db 61 CFTTDPDSIRWEYCNLRK 78
RESULT 3
US-08-763-528A-5
Sequence 5, Application US/08763528A
Patent No. 5854221
GENERAL INFORMATION:
APPLICANT: Cao, Yihai
APPLICANT: Folkman, M. Judah
TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
TITLE OF INVENTION: and Method of Use
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew, LLP
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/763,528A
FILING DATE: 12-DEC-1996
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05940-0251
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Protein
LOCATION: 1.78
OTHER INFORMATION: /note= "Kringle 4 - Figure 3"
US-08-763-528A-5
Query Match 79.7%; Score 417; DB 2; Length 78;
Best Local Similarity 85.9%; Pred. No. 3.1e-40;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
Qy 4 CYHNGQSYRGTSFTVTGTCQSWSSMTPHRRHQRTPEYFNDGLTMNYCRNPDAQTGPW 63
Db 1 CYHGDGQSYRGTSFTVTGTCQSWSSMTPHRRHQRTPEYFNDGLTMNYCRNPDAQTGPW 60
Qy 64 CFTTDPDSIRWEYCNLRK 81
Db 61 CFTTDPDSIRWEYCNLRK 78
RESULT 4
US-09-066-028-23
Sequence 23, Application US/09066028
Patent No. 6024588
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
APPLICANT: O'Reilly, Micheal
APPLICANT: Cao, Yihai
APPLICANT: Sim, B. Kim Lee
TITLE OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/066,028
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/612,788
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799

```

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-066-028-23

Query Match 79.7%; Score 417; DB 3; Length 78;
Best Local Similarity 85.9%; Pred. No. 3.le-40;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTCGTCQSSWTPHRHQRTPENYPNDGLTMNYCRNPADTGPW 63
Db 1 CYHGDGQSYRGTSSTTTGKCKQSSWTPHRHQRTPENYPNAGLTMYCRNPADKGPW 60
QY 64 CFTTDPSSIRWEYCNLRC 81
Db 61 CFTTDPSSIRWEYCNLRC 78

RESULT 5
US-09-335-325-23
; Sequence 23, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Michael
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-08-612-788-22

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-335-325-23

Query Match 79.7%; Score 417; DB 4; Length 78;
Best Local Similarity 85.9%; Pred. No. 3.le-40;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTCGTCQSSWTPHRHQRTPENYPNDGLTMNYCRNPADTGPW 63
Db 1 CYHGDGQSYRGTSSTTTGKCKQSSWTPHRHQRTPENYPNAGLTMYCRNPADKGPW 60
QY 64 CFTTDPSSIRWEYCNLRC 81
Db 61 CFTTDPSSIRWEYCNLRC 78

RESULT 6
US-08-612-788-22
; Sequence 22, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Michael
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-08-612-788-22
```

```
Query Match 70.6%; Score 369; DB 2; Length 78;
Best Local Similarity 73.1%; Pred. No. 8.6e-35;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEYNDGLTMNYCRNPDADTGPW 63
  || :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 1 CYQSDGQSYRGFTSTVTGTCQSWAAMPFPHRSKTPENFPDAGLENNYCRNPDGDKGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CTTTDFSVRWEYCNLKRC 78

RESULT 7
US-09-066-028-22
; Sequence 22, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
US-09-066-028-22

Query Match 70.6%; Score 369; DB 3; Length 78;
Best Local Similarity 73.1%; Pred. No. 8.6e-35;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEYNDGLTMNYCRNPDADTGPW 63
  || :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 1 CYQSDGQSYRGFTSTVTGTCQSWAAMPFPHRSKTPENFPDAGLENNYCRNPDGDKGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CTTTDFSVRWEYCNLKRC 78

RESULT 8
US-09-335-325-22
; Sequence 22, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Sim, B. Kim Lee
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-335-325-22

Query Match 70.6%; Score 369; DB 4; Length 78;
Best Local Similarity 73.1%; Pred. No. 8.6e-35;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQRTPEYNDGLTMNYCRNPDADTGPW 63
  || :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 1 CYQSDGQSYRGFTSTVTGTCQSWAAMPFPHRSKTPENFPDAGLENNYCRNPDGDKGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CTTTDFSVRWEYCNLKRC 78
```



```
;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-335-325-17

Query Match 59.8%; Score 313; DB 4; Length 78;
Best Local Similarity 61.5%; Pred. No. 1.9e-28;
Matches 48; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSPHQRTPENYDGLTMYCRNPADTGPW 63
DB 1 CLKRGENVYRGVSVTGTCQSWSPHQRTPENYDGLTMYCRNPADTGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
DB 61 CYTTSQSLWEYCEIPSC 78

RESULT 12
US-08-612-788-20
; Sequence 20, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
```

```
;
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Porcine
; IMMEDIATE SOURCE:
; CLONE: K3
; US-08-612-788-20

Query Match 57.2%; Score 299; DB 2; Length 78;
Best Local Similarity 56.4%; Pred. No. 7.4e-27;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSPHQRTPENYDGLTMYCRNPADTGPW 63
DB 1 CLKRGENVYRGVSVTGTCQSWSPHQRTPENYDGLTMYCRNPADTGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
DB 61 CYTTSQSLWEYCEIPSC 78

RESULT 13
US-09-066-028-20
; Sequence 20, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
```



```

TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Porcine
IMMEDIATE SOURCE:
CLONE: K3
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-335-325-20

Query Match          57.2%; Score 299; DB 4; Length 78;
Best Local Similarity 56.4%; Pred. No. 7.4e-27;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY      4 CYHNGQSVRGTFSTVTGRTCSWSMSPHPRHQRTPENYDGLTMNYCRNPDAQTGPW 63
      I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Ddb      1 CLKRGENVRGTVSVTASGHCTCQWSAQSPHKNHTPENFFCKNLEENYCRNPDSGTAPW 60
      I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I

QY      64 CFTTDPSTRWECNLTTC 81
      I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Ddb      61 CYTTDSEVRWDYCKIPSC 78
      I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I

RESULT 15
US-08-612-788-18
: Sequence 18, Application US/08612788
: Patent No. 5837682
: GENERAL INFORMATION:
: APPLICANT: Folkman, M. Judah
: APPLICANT: O'Reilly, Micheal
: APPLICANT: Cao, Yihai
: APPLICANT: Sim, B. Kim Lee
: TITLE OF INVENTION: Angiostatin Fragments and Method of Use
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Jones & Askew
: STREET: 191 Peachtree Street, 37th Floor
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: U.S.
: ZIP: 30303-1769
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/612,788
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Warren, William L.
: REGISTRATION NUMBER: 36,714
: REFERENCE/DOCKET NUMBER: 05213-0126
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-818-3700
: TELEFAX: 404-818-3799
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 78 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:35:54 : Search time 61 Seconds
(without alignments)
220.760 Million cell updates/sec

Title: US-10-088-548-6

Perfect score: 523

Sequence: 1 VRQYHGNGSQVRGFTFTV.....SIRWEYCNLTRCSDTEGTGV 89

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 201696

Minimum DB seq length: 0

Maximum DB seq length: 89

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

```
1: /cgn2.6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2.6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2.6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2.6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2.6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2.6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2.6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2.6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2.6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2.6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2.6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2.6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2.6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2.6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2.6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2.6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2.6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2.6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	417	79.7	78	9	US-09-753-064-5
2	417	79.7	78	9	US-09-761-120-23
3	417	79.7	78	10	US-09-335-325-23
4	417	79.7	78	12	US-10-267-137-9
5	417	79.7	78	15	US-10-131-241-23
6	369	70.6	78	9	US-09-761-120-22
7	369	70.6	78	10	US-09-335-325-22
8	369	70.6	78	15	US-10-131-241-22
9	313	59.8	78	9	US-09-761-120-17
10	313	59.8	78	15	US-09-335-325-17
11	313	59.8	78	15	US-10-131-241-17
12	299	57.2	78	9	US-09-761-120-20
13	299	57.2	78	10	US-09-335-325-20
14	299	57.2	78	15	US-10-131-241-20
15	285	54.5	78	10	US-09-335-325-18

16	285	54.5	78	15	US-10-131-241-18
17	283	54.1	78	9	US-09-761-120-19
18	283	54.1	78	10	US-09-335-325-19
19	283	54.1	78	15	US-10-131-241-19
20	281	53.7	78	9	US-09-753-064-4
21	281	53.7	78	9	US-09-761-120-18
22	281	53.7	78	9	US-09-761-120-21
23	281	53.7	78	10	US-09-335-325-21
24	281	53.7	78	15	US-10-131-241-21
25	277	53.0	78	9	US-03-761-120-16
26	277	53.0	78	10	US-09-335-325-16
27	277	53.0	78	15	US-10-131-241-16
28	275	52.6	78	12	US-10-267-137-12
29	274	52.4	78	9	US-09-761-120-15
30	274	52.4	78	10	US-09-335-325-15
31	274	52.4	78	15	US-10-131-241-15
32	268.5	51.3	79	9	US-09-761-120-9
33	268.5	51.3	79	10	US-09-335-325-9
34	268.5	51.3	79	15	US-10-131-241-9
35	267.5	51.1	79	9	US-09-753-064-2
36	267.5	51.1	79	9	US-09-761-120-8
37	267.5	51.1	79	10	US-09-335-325-8
38	267.5	51.1	79	12	US-10-267-137-7
39	267.5	51.1	79	15	US-10-131-241-8
40	264	50.5	78	10	US-09-335-325-13
41	264	50.5	78	15	US-10-131-241-13
42	262	50.1	78	9	US-09-761-120-14
43	262	50.1	78	10	US-09-335-325-14
44	262	50.1	78	15	US-10-131-241-14
45	261	49.9	78	9	US-09-753-064-3

ALIGNMENTS

RESULT 1

```
US-09-753-064-5
; Sequence 5, Application US/09753064
; Patent No. US20010016644A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; Folkman, M. Judah
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
; and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,064
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/763,528
; FILING DATE: 12-Dec-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
```

Sequence 18, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 4, Appl
Sequence 18, Appl
Sequence 21, Appl
Sequence 21, Appl
Sequence 21, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 12, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 2, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 7, Appl
Sequence 8, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 3, Appl

```
;
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..78
; OTHER INFORMATION: /note= "Kringle 4 - Figure 3"
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-753-064-5

Query Match          79.7%; Score 417; DB 9; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSSWSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGWP 63
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||
Db 1 CYHGDGQSYRGFTSTVTGTCQSSWSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGWP 60
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||

QY 64 CFTTDPSSIRWEYCNLTRC 81
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||
Db 61 CFTTDPSSIRWEYCNLKKC 78
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||

RESULT 2
US-09-761-120-23
; Sequence 23, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 23
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 4
US-09-761-120-23

Query Match          79.7%; Score 417; DB 9; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSSWSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGWP 63
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||
Db 1 CYHGDGQSYRGFTSTVTGTCQSSWSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGWP 60
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||

QY 64 CFTTDPSSIRWEYCNLTRC 81
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||
Db 61 CFTTDPSSIRWEYCNLKKC 78
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||

RESULT 3
US-09-335-325-23
; Sequence 23, Application US/09335325
; Patent No. US2002016471A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
```

```
;
; O'Reilly, Micheal
; Cao, Yinai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-335-325-23

Query Match          79.7%; Score 417; DB 10; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSSWSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGWP 63
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||
Db 1 CYHGDGQSYRGFTSTVTGTCQSSWSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGWP 60
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||

QY 64 CFTTDPSSIRWEYCNLTRC 81
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||
Db 61 CFTTDPSSIRWEYCNLKKC 78
   |||||:||||| ||| ||: |||||:|||||:||||| |||||:||||| |||

RESULT 4
US-10-267-137-9
; Sequence 9, Application US/10267137
; Publication No. US20030148950A1
; GENERAL INFORMATION:
; APPLICANT: Xin, Li
; APPLICANT: Li, Zai-Ping
; APPLICANT: Gan, Ren-bao
; APPLICANT: Zhou, Qing-wei
; APPLICANT: Xu, Ren
; TITLE OF INVENTION: KRINGLE DOMAIN 1 OF HUMAN HEPATOCYTE
; TITLE OF INVENTION: GROWTH FACTOR AND USES THEREFOR
```

```
; FILE REFERENCE: 524282000100
; CURRENT APPLICATION NUMBER: US/10/267,137
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/328,329
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence alignment of kringle domains of
; OTHER INFORMATION: plasminogen and HGF
US-10-267-137-9

Query Match          79.7%; Score 417; DB 12; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 63
   |||||:||||| ||| |||:|||||:|||||:|||||:|||||:|||||:||||| |||
Db 1 CHHGDQSYRGTSSTTTGKCKQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADKGPW 60
   |||||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||

QY 64 CFTTDPsirweyCNLRC 81
   |||||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||
Db 61 CFTTDPsirweyCNLKKC 78
   |||||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||

RESULT 5
US-10-131-241-23
; Sequence 23, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-23

Query Match          79.7%; Score 417; DB 15; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 63
   |||||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||
Db 1 CHHGDQSYRGTSSTTTGKCKQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADKGPW 60
   |||||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||

QY 64 CFTTDPsirweyCNLRC 81
   |||||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||
Db 61 CFTTDPsirweyCNLKKC 78
   |||||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||

RESULT 6
US-09-761-120-22
; Sequence 22, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmii
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Kringle 4
US-09-761-120-22

Query Match          70.8%; Score 369; DB 9; Length 78;
Best Local Similarity 73.1%; Pred. No. 1.2e-33;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 63
   |||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||
Db 1 CHHGDQSYRGTSSTTTGKCKQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADKGPW 60
   |||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||

QY 64 CFTTDPsirweyCNLRC 81
   |||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||
Db 61 CFTTDPsirweyCNLKKC 78
   |||:||||| ||| |||:|||||:|||||:|||||:|||||:||||| |||

RESULT 7
US-09-335-325-22
; Sequence 22, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
```

```
/
/
/   LENGTH: 78 amino acids
/   TYPE: amino acid
/   STRANDEDNESS: <Unknown>
/   TOPOLOGY: linear
/   MOLECULE TYPE: protein
/   HYPOTHEICAL: NO
/   ANTI-SENSE: NO
/   FRAGMENT TYPE: N-terminal
/   ORIGINAL SOURCE:
/   ORGANISM: Murine
/   IMMEDIATE SOURCE:
/   CLONE: K4
/   SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-335-325-22

Query Match      70.6%; Score 369; DB 10; Length 78;
Best Local Similarity 73.1%; Pred. No. 1.2e-33;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGRTCSQSSMTPHRHQRTPENYPNDGLTMYCRNPDA DTGPW 63
Db 1 CYQSDGQSYRGFTSTVTGRTCSQSSMTPHRHQRTPENYPNDGLTMYCRNPDA DTGPW 60

QY 64 CFTTDPSPRWECNLKRC 81
Db 61 CYTTDPSRWECNLKRC 78

RESULT 8
US-10-131-241-22
/ Sequence 22, Application US/10131241
/ Publication No. US20030012792A1
/ GENERAL INFORMATION:
/ APPLICANT: Holaday, John W.
/ TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
/ FILE REFERENCE: 05213-0344 43170-271565
/ CURRENT APPLICATION NUMBER: US/10/131,241
/ CURRENT FILING DATE: 2002-07-22
/ PRIOR APPLICATION NUMBER: US 09/413,049
/ PRIOR FILING DATE: 1999-10-06
/ PRIOR APPLICATION NUMBER: US 09/316,802
/ PRIOR FILING DATE: 1999-05-21
/ PRIOR APPLICATION NUMBER: US 60/086,586
/ PRIOR FILING DATE: 1998-05-22
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 22
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Murinae sp.
US-10-131-241-22

Query Match      70.6%; Score 369; DB 15; Length 78;
Best Local Similarity 73.1%; Pred. No. 1.2e-33;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGRTCSQSSMTPHRHQRTPENYPNDGLTMYCRNPDA DTGPW 63
Db 1 CYQSDGQSYRGFTSTVTGRTCSQSSMTPHRHQRTPENYPNDGLTMYCRNPDA DTGPW 60

QY 64 CFTTDPSPRWECNLKRC 81
Db 61 CYTTDPSRWECNLKRC 78

RESULT 9
US-09-761-120-17
/ Sequence 17, Application US/09761120
/ Patent No. US20020037847A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
```

```
/
/
/   APPLICANT: O'Reilly, Michael
/   TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
/   FILE REFERENCE: 05940-0151 (43171-252068)
/   CURRENT APPLICATION NUMBER: US/09/761,120
/   CURRENT FILING DATE: 2001-01-16
/   PRIOR APPLICATION NUMBER: 09/309,821
/   PRIOR FILING DATE: 1999-05-11
/   PRIOR APPLICATION NUMBER: 08/866,735
/   PRIOR FILING DATE: 1997-05-30
/   NUMBER OF SEQ ID NOS: 47
/   SOFTWARE: PatentIn version 3.0
/   SEQ ID NO 17
/   LENGTH: 78
/   TYPE: PRT
/   ORGANISM: Murinae gen. sp.
/   FEATURE:
/   NAME/KEY: misc_feature
/   OTHER INFORMATION: Kringle 3
US-09-761-120-17

Query Match      59.8%; Score 313; DB 9; Length 78;
Best Local Similarity 61.5%; Pred. No. 1.9e-27;
Matches 48; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGRTCSQSSMTPHRHQRTPENYPNDGLTMYCRNPDA DTGPW 63
Db 1 CLKGRGNYRGFTSVTVSGTKCQWSEQPHRHRTPENPKNLEENYCRNPDCGTAPW 60

QY 64 CFTTDPSPRWECNLKRC 81
Db 61 CYTTDPSRWECNLKRC 78

RESULT 10
US-09-335-325-17
/ Sequence 17, Application US/09335325
/ Patent No. US20020164717A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ O'Reilly, Micheal
/ Cao, Yihai
/ Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
```

```
/
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Murine
/ IMMEDIATE SOURCE:
/ CLONE: K3
/ SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-335-325-17

Query Match          59.8%; Score 313; DB 10; Length 78;
Best Local Similarity 61.5%; Pred. No. 1.9e-27;
Matches 48; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CVHNGOSYRGTFSTVTGRTQCSWSTPHRHQRTPENYPNDGLTMNFCRNPDADTGPW 63
DB 1 CLKGGENYRGTVSVTSKTCQWSEQTPHRHRTTENFPCKNLEENYCRNPDEGAPW 60

QY 64 CFTTDPISIRWEYCNLRC 81
DB 61 CYTTDSQLRWEYCEIPSC 78

RESULT 11
US-10-131-241-17
/ Sequence 17, Application US/10131241
/ Publication No. US20030012792A1
/ GENERAL INFORMATION:
/ APPLICANT: Holaday, John W.
/ TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
/ FILE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
/ FILE REFERENCE: 05213-0344 43170-271565
/ CURRENT APPLICATION NUMBER: US/10/131,241
/ CURRENT FILING DATE: 2002-07-22
/ PRIOR APPLICATION NUMBER: US 09/413,049
/ PRIOR FILING DATE: 1999-10-06
/ PRIOR APPLICATION NUMBER: US 09/316,802
/ PRIOR FILING DATE: 1999-05-21
/ PRIOR APPLICATION NUMBER: US 60/086,586
/ PRIOR FILING DATE: 1998-05-22
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 17
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Murine
US-10-131-241-17

Query Match          59.8%; Score 313; DB 15; Length 78;
Best Local Similarity 61.5%; Pred. No. 1.9e-27;
Matches 48; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CVHNGOSYRGTFSTVTGRTQCSWSTPHRHQRTPENYPNDGLTMNFCRNPDADTGPW 63
DB 1 CLKGGENYRGTVSVTSKTCQWSEQTPHRHRTTENFPCKNLEENYCRNPDEGAPW 60

QY 64 CFTTDPISIRWEYCNLRC 81
DB 61 CYTTDSQLRWEYCEIPSC 78

RESULT 12
US-09-761-120-20
/ Sequence 20, Application US/09761120
/ Patent No. US2002037847A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Michael
```

```
/
/ TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
/ FILE REFERENCE: 05940-0151 (43171-252088)
/ CURRENT APPLICATION NUMBER: US/09/761,120
/ CURRENT FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: 09/309,821
/ PRIOR FILING DATE: 1999-05-11
/ PRIOR APPLICATION NUMBER: 08/866,735
/ PRIOR FILING DATE: 1997-05-30
/ NUMBER OF SEQ ID NOS: 47
/ SOFTWARE: Patent in version 3.0
/ SEQ ID NO 20
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Sus sp.
/ FEATURE:
/ NAME/KEY: misc.feature
/ OTHER INFORMATION: Kringle 3
US-09-761-120-20

Query Match          57.2%; Score 299; DB 9; Length 78;
Best Local Similarity 56.4%; Pred. No. 6.9e-26;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 4 CVHNGOSYRGTFSTVTGRTQCSWSTPHRHQRTPENYPNDGLTMNFCRNPDADTGPW 63
DB 1 CLKGGENYRGTVSVTSKTCQWSEQTPHRHRTTENFPCKNLEENYCRNPDEGAPW 60

QY 64 CFTTDPISIRWEYCNLRC 81
DB 61 CYTTDSQLRWEYCEIPSC 78

RESULT 13
US-09-335-325-20
/ Sequence 20, Application US/09335325
/ Patent No. US20020164717A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Michael
/ Cao, Yihai
/ Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
```



```
QY 34 HRHQTPEPNYPNDG-----LTMNYCRNPADTGPWCFTTDPDSIRWEYCNLTRCS 83
Db 5 NQOQSTPEOPTNGQCYIKDCQNKTNWTHRGSRTRDRCGCPKVKPGINLRCKCKDKONE 64

RESULT 3
NLR12
short neurotoxin 2 - ringhals
N;Alternate names: neurotoxin IV
C;Species: Hemachatus haemachatus, Sepedon haemachatus (ringhals)
C;Date: 13-Jul-1991 #sequence_revision 13-Jul-1981 #text_change 23-Aug-1996
C;Accession: A01701
R;Strýdom, A.J.C.; Botes, D.P.
J. Biol. Chem. 246, 1341-1349, 1971
A;Title: Snake venom toxins. Purification, properties, and complete amino acid sequence
A;Reference number: A92073; MUID:71116407; PMID:5545078
A;Accession: A01701
A;Molecule type: protein
A;Residues: 1-61 <SWR>
C;Superfamily: snake toxin
C;Keywords: neurotoxin; venom
F;3-23,17-40,42-53,54-59/Disulfide bonds: #status predicted

Query Match 11.8%; Score 61.5; DB 1; Length 61;
Best Local Similarity 26.1%; Pred. No. 12;
Matches 23; Conservative 8; Mismatches 20; Indels 37; Gaps 5;

QY 3 QCYHGNGQSVRGFTFTVIGTGC-----QSWSSMTHRHQRTPEPNYPNDGLTMNYCR 54
Db 2 ECH--NQOS-----SQTPTQCFGETNCKYKQWSD--HRGSRT----- 36

QY 55 NPADATGCPWCTTDPDSIRWEYCNLTRCS 82
Db 37 -----ERCGGCPVTKPGIKLKCCITDRCN 60

RESULT 4
S24216
I9 kappa chain - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S24216
R;Shimizu, T.; Iwasato, T.; Yamagishi, H.
J. Exp. Med. 173, 1065-1072, 1991
A;Title: Deletions of immunoglobulin C(kappa) region characterized by the circular excis
A;Reference number: S24214; MUID:91217618; PMID:1902500
A;Accession: S24216
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-67 <SH1>
A;Cross-references: EMBL:X58221; NID:g53719; PIDN:CAA41185.1; PID:g930196
C;Superfamily: immunoglobulin V region; Immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 10.3%; Score 54; DB 2; Length 67;
Best Local Similarity 39.5%; Pred. No. 80;
Matches 15; Conservative 1; Mismatches 18; Indels 4; Gaps 1;

QY 1 VRQCYHNGQSVRGFTFS-----TTVTGRTCCQSWSSMTPH 34
Db 20 VRFGSGSGTYSLTISRMEARDATYTCQWSYPPH 57

RESULT 5
T34687
hypothetical protein SC1B5.09c SC1B5.09c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C;Accession: T34687
R;Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1998
A;Reference number: 221553
```

```
A;Accession: T34687
A;Status: preliminary; translated from GS/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-76 <HAR>
A;Cross-references: EMBL:AL023517; PIDN:CAA18983.1; GSPDB:GM00070; SCOEDB:SC1B5.09c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC1B5.09c
```

```
Query Match 9.9%; Score 52; DB 2; Length 76;
Best Local Similarity 27.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 2; Mismatches 24; Indels 28; Gaps 4;
```

```
QY 34 HRHQTPEPNYPNDGLTM---NYC-----RNPADATGPPWC-----FTTD 68
```

```
Db 5 HGDREPE---PDGAGEPGGYCLIDAPRPKADGPFPAECVQCRRPTEYPESYAGITLC 61
```

```
QY 69 PSIRWEYCNLTRCS 82
```

```
Db 62 PVCEWQEAQRTACS 75
```

RESULT 6

A41441

intestinal trefoil factor precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Jul-1999

C;Accession: A41441

R;Suenori, S.; Lynch-Devaney, K.; Podolsky, D.K.

Proc. Natl. Acad. Sci. U.S.A. 88, 11017-11021, 1991

A;Title: Identification and characterization of rat intestinal trefoil factor: tissue

A;Reference number: A41441; MUID:92107881; PMID:1763017

A;Accession: A41441

A;Molecule type: mRNA

A;Residues: 1-81 <SUE>

A;Cross-references: GB:M80826; NID:g207446; PIDN:AAA42270.1; PID:g207447

A;Note: the authors translated the codon ACA for residue 8 as Ile

C;Superfamily: secretory protein xP1; trefoil homology

C;Keywords: intestine

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-22/Domain: propeptide #status predicted <PRO>

F;23-81/Product: intestinal trefoil factor #status predicted <IAT>

F;33-73/Domain: trefoil homology <TRF>

F;33-59,43-58,53-70/Disulfide bonds: #status predicted

```
Query Match 9.9%; Score 52; DB 2; Length 81;
Best Local Similarity 28.2%; Pred. No. 1.5e+02;
Matches 20; Conservative 8; Mismatches 17; Indels 26; Gaps 6;
```

```
QY 16 FSTT-----VTGRTC--QSWSSMTPHR-----HQRTPEPNYPNDGLTMNYCRN----- 55
```

```
Db 6 FWTLLLVLAGSSCKAQEFVGLSPQCMAPTNVRVDCNYPT--VTSEQNNRGCCFDSS 63
```

```
QY 56 PDADTGPWCF 65
```

```
Db 64 IPNV---PWCF 71
```

RESULT 7

B60549

kappa-3 neurotoxin precursor - many-banded krait

N;Alternate names: kappa-3 bungarotoxin; kappa-neurotoxin CRI

C;Species: Bungarus multicinctus (many-banded krait)

C;Date: 17-Apr-1993 #sequence_revision 18-Jun-1993 #text_change 17-Mar-2000

C;Accession: S08398; B60549

R;Danse, J.M.; Garnier, J.M.

Nucleic Acids Res. 18, 1050, 1990

A;Title: cDNA deduced amino-acid sequences of two novel kappa-neurotoxins from Bungar

A;Reference number: S08398; MUID:90192091; PMID:2315018

A;Accession: S08398

A;Molecule type: mRNA

A;Residues: 1-87 <DAN>

A;Cross-references: EMBL:X51413; NID:g62481; PIDN:CAA35775.1; PID:g62482

Query Match 9.5%; Score 49.5; DB 2; Length 88;
Best Local Similarity 29.4%; Pred. No. 3e+02;
Matches 15; Conservative 7; Mismatches 14; Indels 15; Gaps 2;

L₂

RS011 12
S41611
high-potential iron sulfur protein isoform 1 - Ectothiorhodospira vacuolata
N:Alternate names: HiPIP protein

RESULT 15

G69176
 hypothetical protein MTH58 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: G69176
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
 A:Reference number: A59000; NUID:98037514; PMID:9371463
 A:Accession: G69176
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-85 <MTH>
 A:Cross-references: GB:AF000798; GB:AE000666; NID:g2621094; PIDN:AAB84565.1; PID:g262109
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH58

Query Match 9.1%; Score 47.5; DB 2; Length 85;
 Best Local Similarity 25.0%; Pred. No. 4.7e+02;
 Matches 19; Conservative 12; Mismatches 26; Indels 19; Gaps 4;
 QY 18 TTVTGRTQCWSSMTPHRHQTPEH---YPNDGLTMNYCRNPD-ADTGPWCFTTDPISRW 73
 Db 8 TGITCK-----QRVPIDGVYPEDSYTLVETWKPGLFEGP--YTARTEIKY 52
 QY 74 EYCNLTRCSDEGTGV 89
 Db 53 GRFHQTKLETSDIVI 68

Search completed: September 29, 2003, 08:36:31
 Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:23:28 ; Search time 23 Seconds
(without alignments)
181.973 Million cell updates/sec

Title: US-10-088-548-6

Perfect score: 523
Sequence: 1 VRCVHGNGSYRGTFSTTV.....SIRWEYCNLTRCSDEGTIVV 89

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 11264

Minimum DB seq length: 0

Maximum DB seq length: 89

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Match	Length	DB ID	Description
1	64	12.2	64	1	NXS1_BUNFA	P10808 bungarus fa
2	61.5	11.8	61	1	NXS2_HEMHA	P01433 hemachatus
3	50.5	9.7	81	1	POC3_SYRVU	P58171 syringa vul
4	50.5	9.7	84	1	POC3_OLEEU	O81092 olea europa
5	50	9.6	87	1	NXP3_BUNMU	P15817 bungarus mu
6	49	9.4	55	1	ATP8_LATCR	O03168 latimeria c
7	48	9.2	72	1	HP11_ECTVA	P38941 ectothiorho
8	47.5	9.1	84	1	TFP1_HUMAN	P04155 homo sapien
9	47	9.0	60	1	NXS1_ENHSC	P25492 enhydrina s
10	46	8.8	81	1	TFP3_RAT	Q03191 rattus norv
11	45.5	8.7	37	1	SC1T_MBSTA	P01761 mesobuthus
12	45.5	8.7	77	1	THG1_ARATH	Q39182 arabidopsis
13	45	8.6	52	1	YN78_YEAST	P38920 saccharomyc
14	45	8.6	59	1	NAMB_DENJA	P28375 dendroaspis
15	45	8.6	60	1	NXS1_ASTST	P01438 astrotia st
16	45	8.6	60	1	TXW3_NAJHA	P25677 najja haje a
17	45	8.6	61	1	TXWA_NAJHA	P25678 najja haje a
18	45	8.6	70	1	OPX2_RAT	O64201 rattus norv
19	45	8.6	87	1	NXL1_BUNMU	P01398 bungarus mu
20	44.5	8.5	60	1	DISC_ORNMO	P36235 ornithodoros
21	44.5	8.5	62	1	NXS2_LATCO	P10457 laticauda c
22	44.5	8.5	62	1	NKXA_LATCR	P25495 laticauda c
23	44	8.4	61	1	NXS1_HEMHA	P01425 hemachatus
24	44	8.4	88	1	RL34_METHH	O26137 methanobact
25	43.5	8.3	62	1	NXSC_LATCR	P10458 laticauda c
26	43.5	8.3	77	1	THG2_ARATH	Q41914 arabidopsis
27	43.5	8.3	82	1	CXMB_CONNR	Q26443 conus marino
28	43.5	8.3	86	1	TXW2_NAJAT	Q9ygi4 najja atra (
29	43	8.2	35	1	SCX1_BUTSI	P15229 buthus sind
30	43	8.2	60	1	NXS1_AIPLA	P32879 aipysurus l
31	43	8.2	60	1	NXS3_AIPLA	P19958 aipysurus l
32	43	8.2	60	1	SCX3_MESTA	P83400 mesobuthus
33	43	8.2	65	1	YSF1_RHILLO	Q989f2 rhizobium l

ALIGNMENTS

RESULT 1

ID	NXS1_BUNFA	STANDARD;	PRT;	64 AA.
AC	P10808;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Short neurotoxin 1 (Toxin V-II-1).			
OS	Bungarus fasciatus (Banded krait).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;			
OC	Elapidae; Bungarinae; Bungarus.			
OX	NCBI_TaxID=8613;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Venom;			
RX	MEDLINE=89291813; PubMed=2738048;			
RA	Liu C.-S., Chen J.-P., Chang C.-S., Lo T.-B.;			
RT	"Amino acid sequence of a short chain neurotoxin from the venom of banded krait (Bungarus fasciatus).";			
RL	J. Biochem. 105:93-97(1989).			
CC	-!- FUNCTION: PRODUCES PERIPHERAL PARALYSIS BY BLOCKING NEUROMUSCULAR TRANSMISSION AT THE POSTSYNAPTIC SITE. BINDS TO THE NICOTINIC ACETYLCHOLINE RECEPTOR.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- TISSUE SPECIFICITY: Expressed by the venom gland.			
CC	-!- SIMILARITY: Belongs to the snake toxin family.			
DR	PIR: JX0060; JX0060.			
DR	HSP: P01426; LNEA.			
DR	InterPro: IPR003571; Snake_toxin.			
DR	Pfam: PF00087; toxin; 1.			
DR	ProDom: PD000206; Snake_toxin; 1.			
DR	PROSITE; PS00272; SNAKE_TOXIN; 1.			
FT	TOXIN; Neurotoxin; Postsynaptic neurotoxin; Multigene family.			
FT	DISULFID 3 26 BY SIMILARITY.			
FT	DISULFID 20 43 BY SIMILARITY.			
FT	DISULFID 45 56 BY SIMILARITY.			
FT	DISULFID 57 62 BY SIMILARITY.			
SQ	SEQUENCE 64 AA; 7272 MW; 340A6AF40F5E06D5 CRC64;			

Query Match 12.2%; Score 64; DB 1; Length 64;

Best Local Similarity 26.7%; Pred. No. 1.6;

Matches 16; Conservative 7; Mismatches 27; Indels 10; Gaps 1;

QY 34 HEHQTPENYNDG-----LPMVCRNPADTGPWCFTTDSIRWEYCNTRQSD 83

Db 5 NQOQSTPDQPTNQCYIKTKDCKNTWTHRGSRDRCGCPKVKPGINLRCKCTDKNE 64

RESULT 2

ID	NXS2_HEMHA	STANDARD;	PRT;	61 AA.
AC	P01433;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			

Matches	13;	Conservative	6;	Mismatches	21;	Indels	3;	Gaps	1;
----------------	------------	---------------------	-----------	-------------------	------------	---------------	-----------	-------------	-----------

Db 18 FDANGD---GKISSSELGETLKTGLSVTPPEEIORNMAEIDTDG 57

ID	NX13_BUNMU	STANDARD;	PRT	87 AA.
AC	P15817;	P87353;		
DT	01-APR-1990	(Rel. 14, Created		
DT	01-APR-1990	(Rel. 14, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Long neurotoxin CRL precursor (kappa neurotoxin) (Kappa3-bungarotoxin).			
DE	Bungarotoxin).			
OC	Bungarus multicinctus (Many-banded krait).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;			
OC	Elapidae; Bungarinae; Bungarus.			
OX	NCBI_Taxid=8616;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Venom gland;			
EX	MEDLINE=90192891; PubMed=2315018;			
RA	Danse J.M.; Garnier J.M.;			
RT	"CDNA deduced amino-acid sequences of two novel kappa-neurotoxins from Bungarus multicinctus.";			
RL	Nucleic Acids Res. 18:1050-1050(1990).			

RA Chang L.-S., Lin J., Hong E.;
RT "The exon-intron structure of kappa3-bungarotoxin precursor from
RL Taiwan banded krait."; Submitted (MAR-1997) to the EMBL/GenBank/DBD databases.
CC - !- FUNCTION: NEUROTOXIN, BINDS AND INHIBITS NICOTINIC RECEPTORS.
CC COMPARED TO ALPHA-NEUROTOXINS, KAPPA-NEUROTOXIN BIND MORE STRONGLY
CC TO NEURONAL RECEPTORS, AND LESS STRONGLY TO MUSCLE RECEPTORS.
CC - !- SUBCELLULAR LOCATION: Secreted.
CC - !- TISSUE SPECIFICITY: Expressed by the venom gland.
CC - !- SIMILARITY: Belongs to the snake toxin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

[illegible]

Query Match 9.6%; Score 50; DB 1; Length 87;
Best Local Similarity 29.4%; Pred. No. 73;
Matches 20; Conservative 6; Mismatches 24; Indels 18; Gaps 5;

Db

22 RTC-----LISP---SSTPQCPNGDICFRKAQDNFCHSRGVPTEQG--CVATCQCFRS 72

Qy

74 EYCNLPFC 81

Db

73 NYRSLCC 80

```

RESULT 6
ATP8_LATCH
ID      ATP8_LATCH      STANDARD;      PRT;      55 AA.
AC      003168;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN      M1AP8 OR ATP8.
OS      Latimeria chalumnae (Latimeria) (Coelacanth).
OS      Mitochondrion.
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Coelacanthiformes; Coelacanthidae; Latimeria.
OX      NCBI_TaxID=7897;
[1]
RN      SEQUENCE FROM N.A.
RA      Zarogova R., Meyer A.:
RL      Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONMENZIONNED COMPONENT
CC      (CP(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC      -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+){In} = ADP + phosphate +
CC      H(+){Out}.

```

--!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

DR EMBL; U82228; AAC50322.1; -
 DR PIR; E58892; E58892
 DR InterPro; IPR001421; ATPase8_mit.
 DR Pfam; PF00895; ATP-synt-8; 1.
 DR KX Hydrogen ion transport; CF(0);
 TRANSMEM 4
 POTENTIAL; 24
 FO SEQUENCE 55 AA; 6569 MW; 7FE36319E8AF823B CRC54;

Query Match	9.4%;	Score 49;	DB 1;	Length 55;
Best Local Similarity	34.1%;	Pred. No. 56;		
Matches 14;	Conservative	5;	Mismatches	16;
			Indels	6;
			Gaps	3;

RESULT 7	HPIL_ECTVA	HPIL_ECTVA	STANDARD;	PRT;	72 AA.
AD	HPIL_ECTVA				
IC	P38941;				
DT	01-FEB-1995 (Rel. 31, Created)				
DT	01-FEB-1995 (Rel. 31, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	High potential iron-sulfur protein isozyme 1 (HiPIP 1).				
GN	HiPI.				
OS	Ecotothiorhodospira vacuolata.				
OS	Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;				
OC	Ecotothiorhodospiraceae; Ecotothiorhodospira.				
OX	NCBI TaxID=1054;				


```

RN RA
RP RT
RC RL
RX Nucleic Acids Res. 15:1401-1414(1987).
RA "Structure of the human oestrogen-responsive gene p52.";
RX MEDLINE=94145107; PubMed=8311477;
RA Ambler R.P., Meyer T.E., Kamen M.D.;
RT "Amino acid sequences of two high-potential iron sulfur proteins
RT (HIFPs) from the moderately halophilic purple phototrophic bacterium
RT Ectothiorhodospira vacuolata.";
RL Arch. Biochem. Biophys. 308:78-81(1994).
CC -!- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S
CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST
CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST
CC ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA.
CC THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +260 MV.
CC -!- SUBUNIT: Homodimer (Probable).
CC -!- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN
CC (HIFIP) FAMILY.
CC PIR; S41611; S41611.
DR HSSP; P38524; IHPI.
DR InterPro; IPR000170; HipoT_ironulf.
DR Pfam; PF01355; HIFIP; 1.
DR PRINTS; PF00374; HIFIPEROXIN.
DR PROSITE; PS00596; HIFIP; 1.
KW Electron transport; Iron-sulfur; 4Fe-4S.
FT METAL 34 34 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 37 37 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 51 51 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 65 65 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 72 AA; 7690 MW; C5DD79AD593A6F54 CRC64;

Query Match 9.2%; Score 48; DB 1; Length 72;
Best Local Similarity 30.6%; Pred. No. 97;
Matches 15; Conservative 7; Mismatches 15; Indels 12; Gaps 4;

QY 37 QRTPE---NYPNDGLTNYCRNPDATGPGWCT---TDPS-IRWEYC 76
Db 6 ENSPEALALNYKHGASVDH---PSHAQGKNCINCLLYTDPSTFVGWC 51

RESULT 8
TF1FL HUMAN STANDARD; PRT; 84 AA.
AC P04155;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Trefol factor 1 precursor (p52 protein) (HPI.A) (Breast cancer
DE estrogen-inducible protein) (PWR-2).
GN TF1 OR BCE1 OR PS2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=85126393; PubMed=3838275;
RA Prud'Homme J.-F., Fridlansky F., le Cunff M., Atger M.,
RA Mercier-Bodard C., Pichon M.-F., Milgrom E.;
RT "Cloning of a gene expressed in human breast cancer and regulated by
RT estrogen in MCF-7 cells.";
RL DNA 4:11-21(1985).
RN [2]
RX MEDLINE=84169578; PubMed=6324130;
RA Jakowlew S.B., Breathnach R., Jeltsch J.-M., Masiakowski P.,
RA Chambon P.;
RT "Sequence of the p52 mRNA induced by estrogen in the human breast
RT cancer cell line MCF-7.";
RL Nucleic Acids Res. 12:2861-2878(1984).
RN [3]
RX MEDLINE=87146470; PubMed=3822834;
RA Jeltsch J.-M., Roberts M., Schatz C., Garnier J.M., Brown A.M.C.,
RA

```

```

RA Chambon P.;
RT "Structure of the human oestrogen-responsive gene p52.";
RL Nucleic Acids Res. 15:1401-1414(1987).
RN [4]
RX SEQUENCE FROM N.A.
RA MEDLINE=90236983; PubMed=2185238;
RA Mori K., Fujii R., Kida N., Takahashi H., Ohkubo S., Fujino M.,
RA Ohta M., Hayashi K.;
RT "Complete primary structure of the human estrogen-responsive gene
RT (p52) product.";
RL J. Biochem. 107:73-76(1990).
RN [5]
RX SEQUENCE FROM N.A.
RA MEDLINE=90184461; PubMed=2311759;
RA Takahashi H., Kida N., Fujii R., Tanaka K.I., Ohta M., Mori K.,
RA Hayashi K.;
RT "Expression of the p52 gene in human gastric cancer cells derived
RT from poorly differentiated adenocarcinoma.";
RL FEBS Lett. 261:283-286(1990).
RN [6]
RX SEQUENCE FROM N.A.
RA MEDLINE=20408883; PubMed=10950923;
RA Berry A., Scott H.S., Kudoh J., Talior I., Korostishevsky M.,
RA Wattenhofer M., Guipponi M., Barras C., Rossier C., Shibuya K.,
RA Wang J., Kawasaki K., Asakawa S., Minoshima S., Shimizu N.,
RA Antonarakis S., Bonne-Tamir B.;
RT "Refined localization of autosomal recessive nonsyndromic deafness
RT DFNB10 locus using 34 novel microsatellite markers, genomic structure,
RT and exclusion of six known genes in the region.";
RL Genomics 68:22-29(2000).
RN [7]
RX SEQUENCE FROM N.A.
RA MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toroda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal K., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Leirach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [8]
RX SEQUENCE FROM N.A.
RA TISSUE-Colon, Kidney, and Stomach;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallah S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [9]
RX SEQUENCE OF 25-60.

```

RX MEDLINE=68326327; PubMed=3261981;
 RA Mori K., Fujii R., Kida N., Ohta M., Hayashi K.;
 RT "Identification of a polypeptide secreted by human breast cancer
 RL cells (MCF-7) as the human estrogen-responsive gene (ps2) product.";
 RN Biochem. Biophys. Res. Commun. 155:366-372(1988).
 RP
 RX SEQUENCE OF 25-84.
 RA MEDLINE=89119134; PubMed=3146413;
 RT Rio M.-C., Lepage P., Diemunsch P., Roitsch C., Chambon P.;
 RL "Primary structure of human protein ps2.";
 RN C. R. Acad. Sci., III, Sci. Vie 307:825-831(1988).
 RP
 RX BIOSYNTHESIS, AND SECRETION.
 RA MEDLINE=88290699; PubMed=3041593;
 RT Rio M.-C., Bellocq J.P., Daniel J.Y., Tomasetto C., Lathe R.,
 RA Chenard M.P., Batzenschlager A., Chambon P.;
 RT "Breast cancer-associated ps2 protein: synthesis and secretion by
 RL normal stomach mucosa.";
 RN Science 241:705-708(1988).
 RP
 RX STRUCTURE BY NMR.
 RA MEDLINE=96085149; PubMed=8521850;
 RT Polshakov V.I., Frenkiel T.A., Westley B.R., Chadwick M.P.,
 RA May F.E.B., Carr M.D., Feeney J.;
 RT "NMR-based structural studies of the pNR-2/ps2 single domain trefoil
 RL peptide. Similarities to porcine spasmodic peptide and evidence for
 RP a monomeric structure.";
 RN Eur. J. Biochem. 233:847-855(1995).
 RP
 RX STRUCTURE BY NMR.
 RA MEDLINE=97250379; PubMed=9096235;
 RT Polshakov V.I., Williams M.A., Gargaro A.R., Frenkiel T.A.,
 RA Westley B.R., Chadwick M.P., May F.E.B., Feeney J.;
 RT "High-resolution solution structure of human pNR-2/ps2: a single
 RL trefoil motif protein.";
 RN J. Mol. Biol. 267:418-432(1997).
 RP
 RX VARIANTS CANCER ILE-32; ASP-34; LYS-37; ILE-46 AND VAL-55.
 RA MEDLINE=20440596; PubMed=10982763;
 RT Park W.-S., Oh R.-R., Park J.-Y., Lee J.-H., Shin M.-S., Kim H.-S.,
 RA Lee H.-K., Kim Y.-S., Kim S.-Y., Lee S.-H., Yoo N.-J., Lee J.-Y.;
 RT "Somatic mutations of the trefoil factor family 1 gene in gastric
 RL cancer.";
 RN Gastroenterology 119:691-698(2000).
 RP
 RX FUNCTION: Stabilizer of the mucous gel overlying the
 RL gastrointestinal mucosa that provides a physical barrier against
 CC various noxious agents.
 CC
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC
 CC -!- TISSUE SPECIFICITY: FOUND IN STOMACH; IS ALSO EXPRESSED IN BREAST
 CC CANCER BUT NOT IN NORMAL TISSUE. IT IS REGULATED BY ESTROGEN IN
 CC MCF-7 CELLS. STRONG EXPRESSION FOUND IN NORMAL GASTRIC MUCOSA AND
 CC IN THE REGENERATIVE TISSUES SURROUNDING ULCEROUS LESIONS OF
 CC GASTROINTESTINAL TRACT.
 CC
 CC -!- DISEASE: EXPRESSED IN A NUMBER OF CARCINOMAS INCLUDING BREAST
 CC (50%), PANCREAS (70-75%), STOMACH (50-55%), AND LARGE BOWEL (56%).
 CC
 CC -!- SIMILARITY: Contains 1 P-type (trefoil) domain.
 CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/Chromancer/Genes/TF1D201.html".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X00474; CAA25155.1; -;
 CC EMBL; X05030; CAA28695.1; -;
 CC EMBL; X05321; CAA28695.1; JOINED.
 CC EMBL; X05322; CAA28695.1; JOINED.
 CC EMBL; X52003; CAA36254.1; -;
 CC EMBL; M12075; AAA52402.1; -;
 CC

DR EMBL; AB038162; BAB13729.1; -;
 DR EMBL; AP001746; BAA95532.1; -;
 DR EMBL; BC032811; AAH32811.1; -;
 DR PIR; A26667; A26667.
 DR PDB; 1PS2; 07-JUL-97.
 DR PDB; 1HI7; 09-APR-01.
 DR Genew; HGNC:11755; TFF1.
 DR MIM; 113710; -;
 DR GO; GO:0005975; P:carbohydrate metabolism; TAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
 DR InterPro; IPR000519; P:trefoil.
 DR Pfam; PF00088; trefoil.1.
 DR PRINTS; PR00680; PTFREFOIL.
 DR SMART; SM00018; PD; 1.
 DR PROSITE; PS00025; P_TREFOIL; 1.
 DR Growth factor; Signal; Disease mutation; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 84 TREFOIL FACTOR 1.
 FT DOMAIN 30 71 P-TYPE.
 FT DISULFID 31 57 BY SIMILARITY.
 FT DISULFID 41 56 BY SIMILARITY.
 FT DISULFID 51 68 BY SIMILARITY.
 Query Match 9.1%; Score 47.5; DB 1; Length 84;
 Best Local Similarity 31.0%; Pred. No. 1.3e+02;
 Matches 18; Conservative 6; Mismatches 19; Indels 15; Gaps 6;
 Qy 14 GTFTTWTGTCOSWSTPHRPHORTPENYPNDGLTNTYCRN-----PDADTG-PWCF 65
 Db 21 GTLAERQT-ETC-----TVAPERQNC--GFP--GVTPSQCKANGCCFDDTVRGVPWCF 69
 RESULT 9
 NXSL_ENHSC STANDARD; PRT; 60 AA.
 AC P25492; P01436;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Short neurotoxin 1 (Toxin 4).
 OS Enhydrina schistos (Beaked sea snake) (Common sea snake), and
 OS Lepadistia schistos (Hardwick's sea snake).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodactylidae; Squamata; Scleroglossa; Serpentes; Colubroides;
 OC Elapidae; Hydrophiinae; Enhydrina.
 OX NCBI_TaxID=8682, 8781;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=E.schistos; TISSUE=Venom;
 RX MEDLINE=73127342; PubMed=4676045;
 RA Fryklund L., Eaker D., Karlsson E.;
 RT "Amino acid sequences of the two principal neurotoxins of Enhydrina
 RL schistos venom.";
 RL Biochemistry 11:4633-4640(1972).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=L.hardwickii; TISSUE=Venom;
 RX MEDLINE=77246775; PubMed=891966;
 RA Fox J.W., Elzinga M., Tu A.T.;
 RT "Amino acid sequence of a snake neurotoxin from the venom of Lepadistia
 RL hardwickii and the detection of a sulphydryl group by laser Raman
 RL spectroscopy.";
 RL FEBS Lett. 80:217-220(1977).
 CC -!- FUNCTION: PRODUCES PERIPHERAL PARALYSIS BY BLOCKING NEUROMUSCULAR
 CC TRANSMISSION AT THE POSTSYNAPTIC SITE. BINDS TO THE NICOTINIC
 CC ACETYLCHOLINE RECEPTOR.
 CC
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC
 CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC
 CC -!- MISCELLANEOUS: LD(50) IS 0.2 MG/KG BY INTRAVENOUS INJECTION FOR
 CC E. SCHISTOSA.
 CC
 CC -!- SIMILARITY: Belongs to the snake toxin family.
 DR PIR; A01705; NIEY1.
 DR HSSP; P01426; NIEA1.
 DR


```

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: TO YEAST YHR217C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z71614; CAA96274.1; -.
DR EMBL; Z71613; CAA96273.1; -.
DR PIR; S63324; S63324.
DR SGD; S0005282; YNL336W.
KW Hypothetical protein.
SQ SEQUENCE 52 AA; 5951 MW; C1E4066D43E057A1 CRC64;

Query Match      8.6%; Score 45; DB 1; Length 52;
Best Local Similarity 28.9%; Pred. No. 1.4e+02;
Matches 13; Conservative 5; Mismatches 17; Indels 10; Gaps 1;

QY 23 RTQCSWSSM-----TPRHQRTPEPNYNDGLTMYNCRNP 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 RPCLTPSSMQSYDIYPTPTTTHRTHTPTTPHPPHPTHHTHNP 48

RESULT 14
NAME_DENJA
ID NAME_DENJA STANDARD; PRT; 59 AA.
AC P28375;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mambin (Glycoprotein IIB-IIIa antagonist) (Platelet aggregation
DE inhibitor) (dendroaspin).
OS Dendroaspis jamestoni kamosae (Eastern Jameson's mamba).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Dendroaspis.
OX NCBI_TaxID=8619;
RN [1]
RP TISSUE=Venom;
RC MEDLINE=92273586; PubMed=1591238;
RA McDowell R.S., Dennis M.S., Louie A., Shuster M., Mulkerrin M.G.,
RA Lazarus R.A.;
RT "Mambin, a potent glycoprotein IIB-IIIa antagonist and platelet
RT aggregation inhibitor structurally related to the short
RT neurotoxins."
RL Biochemistry 31:4766-4772(1992).
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE=95360732; PubMed=7634091;
RA Sutcliffe M.J., Jaseja M., Hyde E.I., Lu X., Williams J.A.;
RT "Three-dimensional structure of the RGD-containing neurotoxin
RT homologue dendroaspin."
RL Nat. Struct. Biol. 1:802-812(1994).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=95112851; PubMed=7813476;
RA Jaseja M., Lu X., Williams J.A., Sutcliffe M.J., Kakkar V.V.,
RA Parslow R.A., Hyde E.I.;
RT "1H-NMR assignments and secondary structure of dendroaspin, an RGD-
RT containing glycoprotein IIB-IIIa (alpha IIB-beta 3) antagonist with a
RT neurotoxin fold."
RL Eur. J. Biochem. 226:861-869(1994).
CC -I- FUNCTION: INHIBITS ADP-INDUCED PLATELET AGGREGATION AND INHIBITS
CC THE BINDING OF PURIFIED PLATELET FIBRINOGEN RECEPTOR GPIIb-IIIa
CC TO IMMOBILIZED FIBRINOGEN.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -I- SIMILARITY: Belongs to the snake toxin family.

```

```

DR PIR; A42561; T6EP5J.
DR PDB; 1DRS; 20-DEC-94.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR ProDom; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; FALSE_NEG.
KW Blood coagulation; Cell adhesion; Toxin; 3D-structure.
FT DISULFID 3 22
FT DISULFID 17 37
FT DISULFID 39 51
FT DISULFID 52 57
FT SITE 43 45 CELL ATTACHMENT SITE (POTENTIAL).
FT STRAND 2 4
FT TURN 8 9
FT STRAND 14 16
FT STRAND 22 24
FT TURN 28 31
FT STRAND 36 37
FT STRAND 50 52
SQ SEQUENCE 59 AA; 6754 MW; 20DDC6A5D9DF9E41 CRC64;

Query Match      8.6%; Score 45; DB 1; Length 59;
Best Local Similarity 28.2%; Pred. No. 1.6e+02;
Matches 22; Conservative 8; Mismatches 28; Indels 20; Gaps 5;

QY 2 ROCYHGQSYRGTEFTTGTCTQCSWSSMTPRHQRTPEPNYNDGLTMYNCRNP 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 RICYNHLG-----TKPPTTETCQEDSC---YKNWTFDNIIRGCG---CFTPRGDMP 47

QY 61 GPWCFTTDFSIWYCNL 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 GPYCCESDR-----CNL 59

RESULT 15
NAME_ASSTST
ID NAME_ASSTST STANDARD; PRT; 60 AA.
AC P04438; P10461; P19005;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Short neurotoxin 1 (Toxin A).
OS Astrotia stokesi (Stokes's sea snake) (Disteira stokesi),
OS Hydrophis ornatus (Reef sea snake), and
OS Acalyptophis peronii (Sea snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Hydrophiinae; Astrotia.
OX NCBI_TaxID=8680, 8685, 8676;
RN [1]
RP SEQUENCE.
RC SPECIES=A.stokesi; TISSUE=Venom;
RX MEDLINE=79123997; PubMed=743209;
RA Maeda N., Tamiva N.;
RT "Three neurotoxins from the venom of a sea snake Astrotia stokesii,
RT including two long-chain neurotoxic proteins with amidated
RT C-termini."
RL Biochem. J. 175:507-517(1978).
RN [2]
RP COMPOSITION OF TRYPTIC PEPTIDES.
RC SPECIES=H.ornatus; TISSUE=Venom;
RX MEDLINE=83308533; PubMed=6615431;
RA Tamiya N., Maeda N., Cogger H.G.;
RT "Neurotoxins from the venoms of the sea snakes Hydrophis ornatus and
RT Hydrophis lapemoides."
RL Biochem. J. 213:31-38(1983).
RN [3]
RP SEQUENCE (MAJOR VARIANT).
RC SPECIES=A.peronii; TISSUE=Venom;
RX MEDLINE=88132805; PubMed=3341735;
RA Mori N., Tu A.T.;
RT "Isolation and primary structure of the major toxin from sea snake,
RT Acalyptophis peronii, venom."

```

```

RL Arch. Biochem. Biophys. 260:10-17(1988).
RN [4]
RP SEQUENCE (MINOR VARIANT).
RC SPECIES=A. peronii; TISSUE=Venom;
RA MEDLINE=89076487; PubMed=3202959;
RA Mori N., Tu A.T.;
RT "amino-acid sequence of the minor neurotoxin from Acalyptophis
RT peronii venom.";
RL Biol. Chem. Hoppe-Seyler 369:521-526(1988).
CC -!- FUNCTION: PRODUCES PERIPHERAL PARALYSIS BY BLOCKING NEUROMUSCULAR
CC TRANSMISSION AT THE POSTSYNAPTIC SITE. BINDS TO THE NICOTINIC
CC ACETYLCHOLINE RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- MISCELLANEOUS: LD(50) IS 0.170 MG/KG BY INTRAVENOUS INJECTION FOR
CC A. PERONII MINOR TOXIN, 0.125 MG/KG FOR THE MAJOR TOXIN.
CC -!- MISCELLANEOUS: LD(50) IS 0.13 MG/KG BY INTRAMUSCULAR INJECTION FOR
CC A. STOKESI.
CC -!- SIMILARITY: Belongs to the snake toxin family.
DR PIR; A01706; N1AT1.
DR PIR; A27580; A27580.
DR PIR; B90321; N1AT1F.
DR HSSP; P01426; 1NEA.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR ProDom; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Toxin; Neurotoxin; Postsynaptic neurotoxin; Multigene family.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 17 39 BY SIMILARITY.
FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 53 58 BY SIMILARITY.
FT VARIANT 43 43 Q->E (IN A. PERONII MINOR VARIANT).
SQ SEQUENCE 60 AA; 6599 MW; DE41989F90565E9A CRC64;

Query Match 8.6%; Score 45; DB 1; Length 60;
Best Local Similarity 21.8%; Pred. NO. 1.7e+02;
Matches 17; Conservative 12; Mismatches 25; Indels 24; Gaps 4;

Qy 8 NCOSYRGTFSTVTGRC--QSWSSMTFHRHQRTPEYFNDGLTMYCNPADATGPWCF 65
Db 5 NQSSSQPKTTNCAGNSCYKTKWSD-----HRTG-----IIRGCGCPQVKSG---- 47

Qy 66 TTDPSIRWEXCNLRGSD 83
Db 48 -----IKLECCHTNECNN 60

```

Search completed: September 29, 2003, 08:34:03
 Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:30:34 ; Search time 94 Seconds
(without alignments)
244.326 Million cell updates/sec

Title: US-10-088-548-6
Perfect score: 523
Sequence: 1 VRQYHGNGSGYRGTFSTTV.....SIRWEYCNLTRCSDTEGTVV 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 113503

Minimum DB seq length: 0
Maximum DB seq length: 89

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	223	42.6	53	Q9UMI2	Q9um12 homo sapien
2	190	36.3	60	Q9UKJ7	Q9ukj7 homo sapien
3	141	27.0	25	Q9UD88	Q9ud88 homo sapien
4	127.5	24.4	75	Q9BGN9	Q9bgn9 bos taurus
5	59.5	11.4	86	Q9ORV4	Q9orv4 human immun
6	56	10.7	87	Q9OBV2	Q9obv2 human immun
7	55.5	10.6	87	Q9H444	Q9h444 oryza sativ
8	53.5	10.2	85	Q8FTS2	Q8fts2 corynebacte
9	52.5	10.0	82	Q9YL18	Q9yl18 human immun
10	52.5	10.0	87	Q8V5I3	Q8v5i3 human respi
11	52.5	10.0	88	Q9ID04	Q9id04 human immun
12	52.5	10.0	89	Q9ID03	Q9id03 human immun
13	52	9.9	39	Q9MZ7	Q9mf7 bos taurus
14	52	9.9	46	Q9FWJ8	Q9fwj8 gallus gall
15	52	9.9	67	Q9QJH2	Q9qjh2 human immun
16	52	9.9	76	Q9837	Q9837 streptomyce

17	51.5	9.8	72	12	Q91ND7	Q91nd7 cactus viru
18	51.5	9.8	87	12	Q8BBA5	Q8bba5 human respi
19	51.5	9.8	87	12	Q8BBA4	Q8bba4 human respi
20	51	9.8	66	15	Q9QJ14	Q9qj14 human immun
21	51	9.8	83	12	Q8V5H6	Q8v5h6 human respi
22	51	9.8	88	15	Q9DXN2	Q9dxn2 human immun
23	50.5	9.7	57	10	Q8H3H3	Q8h3h3 oryza sativ
24	50.5	9.7	80	12	Q8V5I6	Q8v5i6 human respi
25	50.5	9.7	83	12	Q8BBA3	Q8bba3 human respi
26	50.5	9.7	87	12	Q8V5I7	Q8v5i7 human respi
27	50	9.6	73	15	Q72408	Q72408 human immun
28	50	9.6	83	12	Q9IH88	Q9ih88 human respi
29	50	9.6	83	12	Q8V5I2	Q8v5i2 human respi
30	50	9.6	83	12	Q8BBA8	Q8bba8 human respi
31	50	9.6	84	15	Q9IG35	Q9ig35 human immun
32	50	9.6	88	15	Q9DXK9	Q9dxk9 human immun
33	49.5	9.5	62	5	Q8T0X4	Q8t0x4 mesobuthus
34	49.5	9.5	65	2	Q9E2G6	Q9ezg6 streptococc
35	49.5	9.5	69	2	Q9Z3A6	Q9z3a6 unidentified
36	49.5	9.5	71	15	Q76655	Q76655 human immun
37	49.5	9.5	72	5	Q8MQK9	Q8mqk9 schistosoma
38	49.5	9.5	76	12	Q8VAD4	Q8vad4 white spot
39	49.5	9.5	77	10	Q39403	Q39403 brassica ca
40	49.5	9.5	79	5	O44062	O44062 epicomplexa
41	49.5	9.5	80	12	Q9IUD1	Q9iud1 human respi
42	49.5	9.5	83	12	Q9IUE3	Q9iue3 human respi
43	49.5	9.5	83	12	Q8V5H5	Q8v5h5 human respi
44	49.5	9.5	87	12	Q8V5I4	Q8v5i4 human respi
45	49.5	9.5	87	15	Q90A68	Q90a68 human immun

ALIGNMENTS

RESULT 1
Q9UMI2 ID Q9UMI2 PRELIMINARY; PRT; 53 AA.
AC Q9UMI2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE PLG protein (fragment).
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for human and bovine plasminogen."
RL Biochemistry 23:4243-4250(1984).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; K02921; AAA60123.1; -;
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle1.
DR PRINTS; PRO0018; KRINGLE.
DR PRODOM; PD000395; Kringle.1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
FT NON_TER 1
SQ SEQUENCE 53 AA; 5893 MW; 2EE36E715CD40A47 CRC64;

Query Match 42.6%; Score 223; DB 4; Length 53;
Best Local Similarity 82.2%; Pred. No. 1.1e-18;
Matches 37; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VRQYHGNGSGYRGTFSTTVGTTCQSWSMTPHRHQTPEYVPN 45


```

RA Sullivan M., Gwinn M., Clark K.A., Busch M.P.;
RT "Use of the Dual Sensitive/Less-Sensitive (Detuned) EIA Strategy for
RT Targeting Genetic Analysis of HIV-1 to Recently Infected U.S. Blood
RT Donors.;"
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF373767; AAK53387.1; -
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 86
SQ SEQUENCE 86 AA; 9895 MW; 1A0C197C8C58A808 CRC64;

Query Match 11.4%; Score 59.5; DB 15; Length 86;
Best Local Similarity 30.0%; Pred. No. 20;
Matches 21; Conservative 10; Mismatches 18; Indels 21; Gaps 4;

QY 41 ENYPNDGLTM-----NYCRNPADT-----GFWCFTTD--PSIRWEYCN--LT 79
||: : : : ||: : : : ||: : : : ||: : : : ||: : : : ||
Db 10 ENFYDNAKTIILQNSGVICERPNNTRKGIHIGKVFYTTDIIIGDKAKCNISLT 69
||: : : : ||: : : : ||: : : : ||: : : : ||: : : : ||
QY 80 RCDSTEGTVV 89
: : || : :
Db 70 KWNDTLKQIV 79

RESULT 6
Q90BU2 PRELIMINARY; PRT; 87 AA.
AC Q90BU2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P64-49;
RA Collins K., Quinones-Mateu M., Wu M., Luze H., Johnson J., Hirsch C.,
RA Toossi Z., Arts E.;
RT "A TB-mediated site Specific Increase in HIV-1 Quasispecies Diversity
RT May Lead to the Observed Greater HIV-1 Heterogeneity in Patients Co-
RT Infected with Tuberculosis.;"
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF383521; AAK98372.1; -
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9930 MW; 1D246B516BDD1F2C CRC64;

Query Match 10.7%; Score 56; DB 15; Length 87;
Best Local Similarity 32.7%; Pred. No. 51;
Matches 17; Conservative 9; Mismatches 14; Indels 12; Gaps 4;

QY 45 NDGLTMNYCRNPADT-----GP---WCFTTD--PSIRWEYCNTRCSDTE 85
||: : : : ||: : : : ||: : : : ||: : : : ||: : : : ||
Db 16 NESVQIN-CTRPNNTKSVHIGPGQAFYATGDIIGKIRWAYCNYSRAAWNE 66
||: : : : ||: : : : ||: : : : ||: : : : ||: : : : ||

RESULT 7
Q8H444 PRELIMINARY; PRT; 87 AA.
AC Q8H444;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P0407H12.22 protein.
GN P0407H12.22.

OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone: P0407H12.;"
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP004303; BAC21447.1; -
SQ SEQUENCE 87 AA; 10093 MW; 5BC46C62C68F3BF CRC64;

Query Match 10.6%; Score 55.5; DB 10; Length 87;
Best Local Similarity 30.5%; Pred. No. 58;
Matches 18; Conservative 3; Mismatches 29; Indels 9; Gaps 3;

QY 35 RHQRTPEN-YPNDGLTMNYCRNPADTGPWCFTTDSIRWEYCNLT---RCDSTEGTVV 89
||| : : : : ||: : : : ||: : : : ||: : : : ||: : : : ||
Db 12 RHQAWPATAMEDDGRKDDAAAPGDSHRW-----PQIRWRPCQPPRHHCSTPPGRAV 65
||| : : : : ||: : : : ||: : : : ||: : : : ||: : : : ||

RESULT 8
Q8FTS2 PRELIMINARY; PRT; 85 AA.
AC Q8FTS2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN CEL489.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.;"
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005219; BAC18299.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 85 AA; 9868 MW; AA74412387441A4A CRC64;

Query Match 10.2%; Score 53.5; DB 16; Length 85;
Best Local Similarity 27.0%; Pred. No. 97;
Matches 17; Conservative 6; Mismatches 19; Indels 21; Gaps 4;

QY 28 WSSMTPH-----RHQRTPE-----NYPNDGLTMNYCRNPADTGPWCFTTDSIRWE 74
||| : : : : ||: : : : ||: : : : ||: : : : ||: : : : ||
Db 10 WGSPLPEILGGARSYLITEGSPADPGLENSKIRVNSCHR-----WCFKREISIR-D 61
||: : : : ||: : : : ||: : : : ||: : : : ||: : : : ||
QY 75 YCN 77
||:
Db 62 YCS 64

RESULT 9
Q9YL18 PRELIMINARY; PRT; 82 AA.
AC Q9YL18;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

```

OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UG5609;
RA MEDLINE=9910134; PubMed=10195759;
RA Naghavi M.H., Salminen M.O., Sonnerborg A., Vahlne A.;
RT "DNA sequence of the long terminal repeat of human Immunodeficiency
RT virus type 1 subtype A through G.";
RL AIDS Res. Hum. Retroviruses 15:485-488(1999).
DR EMBL: AF106153; AAD16901.1;
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9408 MW; 731C8B5B2D075638 CRC64;

Query Match 10.08; Score 52.5; DB 15; Length 82;
Best Local Similarity 34.8%; Pred. No. 1.3e+02;
Matches 16; Conservative 7; Mismatches 12; Indels 11; Gaps 3;

Oy 45 NDGLTMVCRNPADT-----GWCFTTD--PSIRWEYCNLTR 80
Db 13 NESVIN-CTRPYNTQSIHMGPRAYTTDITGDIRAYCNISR 57

RESULT 10
OQ8V5I3
ID Q8V5I3 PRELIMINARY; PRT; 87 AA.
AC Q8V5I3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Attachment glycoprotein G (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA98V153;
RX MEDLINE=21405813; PubMed=11514720;
RA Venter M., Madhi S.A., Tiemessen C.T., Schoub B.D.;
RT "Genetic diversity and molecular epidemiology of respiratory syncytial
RT virus over four consecutive seasons in South Africa: identification of
RT new subgroup A and B genotypes.";
RL J. Gen. Virol. 82:2117-2124(2001).
DR EMBL: AF348815; AAL60316.1;
DR InterPro: IPR000925; Glycoprot_G.
DR Pfam: PF00802; Glycoprotein_G; 1.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9431 MW; CEF8B4DD27344769 CRC64;

Query Match 10.08; Score 52.5; DB 12; Length 87;
Best Local Similarity 34.9%; Pred. No. 1.3e+02;
Matches 15; Conservative 1; Mismatches 22; Indels 5; Gaps 1;

Oy 7 GNGQSYRGFTTFTVTGTCQSWSPPHRQHTPEYTPNDGLT 49
Db 31 GTSTQSTVLDTTTSKTIQQS-----LHSTTPENTPNSQT 68

RESULT 11
OQ9ID04
ID Q9ID04 PRELIMINARY; PRT; 88 AA.
AC Q9ID04;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Envelope protein (Fragment).
GN ENV OR GP120.
OS Human immunodeficiency virus 1.

```

```

OX Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL692;
RA MEDLINE=99388930; PubMed=10461834;
RA Navas M.C., Letourneur E., Gomas E., Boshell J., Saragosti S.;
RT "Analysis of the V3 loop sequences from 12 HIV type-1 infected
RT patients from Colombia, South America.";
RL AIDS Res. Hum. Retroviruses 15:1141-1144(1999).
DR EMBL: Y10364; CAA71393.1;
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 88
SQ SEQUENCE 88 AA; 9884 MW; 88ADA051D4AD1569 CRC64;

Query Match 10.08; Score 52.5; DB 15; Length 88;
Best Local Similarity 32.6%; Pred. No. 1.3e+02;
Matches 15; Conservative 8; Mismatches 12; Indels 11; Gaps 3;

Oy 45 NDGLTMVCRNPADT-----GWCFTTD--PSIRWEYCNLTR 80
Db 13 NESVIN-CTRPNNTRSRTHVPGKALYTTDTITGDIRAYCNISR 57

RESULT 12
OQ9ID03
ID Q9ID03 PRELIMINARY; PRT; 89 AA.
AC Q9ID03;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope protein (Fragment).
GN ENV OR GP120.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL694;
RX MEDLINE=99388930; PubMed=10461834;
RA Navas M.C., Letourneur E., Gomas E., Boshell J., Saragosti S.;
RT "Analysis of the V3 loop sequences from 12 HIV type-1 infected
RT patients from Colombia, South America.";
RL AIDS Res. Hum. Retroviruses 15:1141-1144(1999).
DR EMBL: Y10366; CAA71395.1;
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 89
SQ SEQUENCE 89 AA; 9858 MW; 688C7D9A96FFE10F CRC64;

Query Match 10.08; Score 52.5; DB 15; Length 89;
Best Local Similarity 31.6%; Pred. No. 1.3e+02;
Matches 18; Conservative 10; Mismatches 16; Indels 13; Gaps 4;

Oy 45 NDGLTMVCRNPADT-----GWCFTTD--PSIRWEYCNLTR 89
Db 13 NESVIN-CTRPNNTRQGVHMGPGKALYTKYGGIRQAYCNVSRWNLTGQIV 68

RESULT 13
OQ9MZF7
ID Q9MZF7 PRELIMINARY; PRT; 39 AA.
AC Q9MZF7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue plasminogen activator (Fragment).
GN PLAT.

```

Search completed: September 29, 2003, 08:35:46
Job time : 97 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:36:39 ; Search time 82 Seconds
(without alignments)
166.469 Million cell updates/sec

Title: US-10-088-548-8

Perfect score: 508

Sequence: 1 EQDCMGNGKGRGKKATTV.....YTMNPKLEDYCDIPLCASS 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 686138

Minimum DB seq length: 0

Maximum DB seq length: 86

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	508	100.0	86	22 AAY72947	Human apolipoprotein
2	413	81.3	85	21 AAB08413	Amino acid sequence
3	405	79.7	79	18 AAW19256	Human plasminogen
4	274.5	54.0	79	17 AAW07546	Rhesus kringle 1.
5	258.5	50.9	79	17 AAW07548	Bovine kringle 1.
6	258.5	50.9	82	17 AAY77719	Human plasminogen
7	258.5	50.9	84	10 AAP93230	Plasminogen kringle
8	258.5	50.9	84	21 AAB01906	Human plasminogen
9	256.5	50.5	79	17 AAW07545	Human kringle 1.

10	254	50.0	81	10 AAP90598	Synthetic N-termin
11	252.5	49.7	84	21 AAB08410	Amino acid sequence
12	251.5	49.5	79	9 AAP81360	K1 domain of plas
13	251.5	49.5	79	17 AAR96221	Plasminogen kringle
14	248.5	48.9	78	10 AAP94401	Sequence encoding
15	246.5	48.5	79	17 AAW07547	Porcine kringle 1.
16	241.5	47.5	79	17 AAW07544	murine kringle 1.
17	240	47.2	78	17 AAW07559	Human kringle 4.
18	238.5	46.9	86	23 ABP02801	Human ORFX protein
19	238	46.9	82	10 AAP93231	Plasminogen kringle
20	235	46.3	84	21 AAY77720	Human plasminogen
21	223	43.9	85	21 AAB01907	Human plasminogen
22	222	43.7	78	17 AAW07554	Murine kringle 3.
23	222	43.7	78	17 AAW07560	Human kringle 4.
24	220	43.3	78	17 AAW07553	Bovine kringle 2.
25	220	43.3	78	17 AAW07557	Porcine kringle 3.
26	217	42.7	78	17 AAW07549	Murine kringle 2.
27	217	42.7	83	21 AAB01908	Human plasminogen
28	216	42.5	78	17 AAW07550	Human kringle 2.
29	216	42.5	78	17 AAW07551	Rhesus kringle 2.
30	215	42.3	78	17 AAW07552	Human kringle 3.
31	215	42.3	78	17 AAW07552	Porcine kringle 2.
32	215	42.3	78	17 AAW07556	Rhesus kringle 3.
33	213	41.9	78	17 AAW07558	Bovine kringle 3.
34	213	41.9	78	21 AAY77721	Human plasminogen
35	213	41.9	83	21 AAB08414	Amino acid sequence
36	212	41.7	83	21 AAB08412	Amino acid sequence
37	194	38.2	82	21 AAB08411	Amino acid sequence
38	149	29.3	47	13 AAR25682	Exon V of human he
39	146	28.7	32	23 AAR98485	Plasminogen peptid
40	136	26.8	30	23 AAY98483	Plasminogen peptid
41	135	26.6	38	21 AAY77722	Human plasminogen
42	131.5	25.9	44	13 AAR25688	Exon XI of human h
43	130.5	25.7	72	22 AAB97380	Human kringle doma
44	128.5	25.3	42	22 ABB48942	Human liver peptid
45	128.5	25.3	42	22 ABB28941	Peptide #1592 enco

ALIGNMENTS

RESULT 1
AAY72947

ID AAY72947 standard; Protein; 86 AA.

XX AAY72947;

AC AAY72947;

XX 13-JUN-2001 (first entry)

DT Human apolipoprotein(a) Kringle domain V38, LK8 protein.

DE Human; angiogenesis inhibitor; LK8; apolipoprotein(a) kringle domain;
KW angiogenesis-mediated disease; cancer; rheumatoid arthritis; therapy;
KW cytostatic; antirheumatic; antiarthritic; antipsoriasis; psoriasis;
KW ocular angiogenic disease; endothelial cell proliferation; tumour;
KW cell migration.

XX Homo sapiens.

XX WO200119868-A1.

XX 22-MAR-2001.

XX 15-SEP-1999; 99WO-KR00554.

XX 15-SEP-1999; 99WO-KR00554.

XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.

XX Chang J, Kim JS, Park EJ, Yum J, Chung S;

XX WPI; 2001-244787/25.

DR N-PSDB; AAD03258.

XX PT Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
PT diseases, e.g. cancer and rheumatoid arthritis, has human
PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
PT sequence -
XX
XX PS Claim 3; Page 47; 50pp; English.
XX CC The present sequence is human LK8 protein which contains the amino
CC acid sequence of human apolipoprotein(a) kringle domain V38
CC (KV38). The human apolipoprotein(a) kringle domains IV36 (LK6
CC protein), IV37 (LK7 protein) and V38 (LK8 protein) together form the
CC angiogenesis inhibitor, LK68 protein. LK68, LK6, LK7 and LK8 are
CC inhibitors of endothelial cell proliferation, cell migration and
CC normal development of capillaries in the chick embryo chorioallantoic
CC membrane (CAM). LK68 protein, its single kringle or their functional
CC equivalents, are useful for treating angiogenesis-mediated diseases,
CC such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic
CC disease in animals or humans. LK68 is useful as an anticancer agent and
CC also for inhibiting primary tumour growth.
XX
XX SQ Sequence 86 AA;
Query Match 100.0%; Score 508; DB 22; Length 86;
Best Local Similarity 100.0%; Pred. No. 1e-42;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGD 60
Db 1 EDCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGD 60
QY 61 INGPWCYTMNPKLFDYCDIPLCASS 86
Db 61 INGPWCYTMNPKLFDYCDIPLCASS 86

RESULT 2
AAB08413
ID AAB08413 standard; Protein; 85 AA.
XX AC AAB08413;
XX
XX DT 20-DEC-2000 (first entry)
XX DE Amino acid sequence of kringle 5 of human plasminogen.
XX KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;
KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;
KW adult respiratory distress syndrome; Castlemann's disease; psoriasis;
KW hepatitis; aneurysm; renal disease; haemangioma.
XX OS Homo sapiens.
XX WO200049871-A1.
XX PN 31-AUG-2000.
XX PD 24-FEB-2000; 2000WO-US04798.
XX PF 24-FEB-1999; 99US-0121341.
XX PR 25-FEB-1999; 99US-0121633.
XX PR 18-NOV-1999; 99US-0166176.
XX
XX (FORD-) FORD HEALTH SYSTEM HENRY.
XX
XX Dou D, Chopp M, Wang L, Mikkelsen T;
XX WPI; 2000-572016/53.
XX
XX Use of kringle protein and kringle derived from plasminogen and
PT composition comprising kringle proteins for treating tumor and
PT atherosclerosis, arthritis and retinopathy -

XX PS Disclosure; Fig 6; 163pp; English.
XX CC The specification describes a human polypeptide which is a potent
CC angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle
CC proteins, or a kringle derived from human tissue plasminogen activator
CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
CC or tPA are useful for treating tumours, as well as atherosclerosis,
CC arthritis, retinopathy and other similar diseases. KED is also useful
CC for the treatment of diseases such as bronchial vascular congestion,
CC inflammatory bowel disease, adult respiratory distress syndrome,
CC Castlemann's disease, psoriasis, hepatitis, aneurysm, renal disease
CC and haemangioma. The present sequence represents kringle 5 of human
CC plasminogen, which is used in the course of the invention.
XX
XX SQ Sequence 85 AA;
Query Match 81.3%; Score 413; DB 21; Length 85;
Best Local Similarity 80.0%; Pred. No. 2.3e-33;
Matches 68; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
QY 1 EDCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGD 60
Db 1 EDCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGD 60
QY 61 INGPWCYTMNPKLFDYCDIPLCASS 85
Db 61 VGGPWCYTTNDRKLYDYCDVPCAA 85
RESULT 3
AAW19256
ID AAW19256 standard; Peptide; 79 AA.
XX AC AAW19256;
XX DT 27-FEB-1998 (first entry)
XX DE Human plasminogen Kringle 5 fragment.
XX KW Plasminogen; Kringle 5; cell proliferation inhibitor; angiogenesis;
KW diagnosis; therapeutic.
XX OS Homo sapiens.
XX PN WO9723500-A1.
XX PD 03-JUL-1997.
XX PF 13-DEC-1996; 96WO-US20447.
XX PR 12-DEC-1996; 96US-0763528.
XX PR 13-DEC-1995; 95US-0008519.
XX PA (CHIL-) CHILDRENS MEDICAL CENT.
XX PI Cao Y, Folkman MJ;
XX WPI; 1997-350965/32.
XX Plasminogen Kringle 5 peptide - which inhibits endothelial cell
PT proliferation, useful to treat angiogenesis mediated diseases and in
PT detection and diagnosis
XX
XX Claim 1; Page 8; 51pp; English.
XX This sequence is an isolated fragment of the Kringle 5 peptide
CC corresponding to amino acid 462 of the human plasminogen protein which
CC can be used in a novel method to inhibit endothelial cell proliferation
CC activity. The protein can be used to treat angiogenesis mediated
CC diseases, e.g. haemangioma, solid tumours, leukaemia, metastasis,
CC telangiectasia, psoriasis, scleroderma, pyogenic granuloma, myocardial
CC angiogenesis, plaque neovascularisation, coronary or cerebral

CC collaterals, arteriovenous malformations, ischemic limb angiogenesis,
 CC corneal diseases, rubeosis, neovascular glaucoma, diabetic retinopathy,
 CC retrolental fibroplasia, arthritis, diabetic neovascularisation,
 CC muscular degeneration, peptic ulcer, Helicobacter related disease,
 CC fractures, keloids, vasculogenesis, haematopoiesis, ovulation,
 CC menstruation, placentaion or cat scratch fever, and to stimulate wound
 CC healing. The protein and antibodies generated from it can be used to
 CC screen for agonists and antagonists or in detection, imaging and
 CC diagnosis.
 XX Sequence 79 AA;
 SQ

Query Match 79.7%; Score 405; DB 18; Length 79;
 Best Local Similarity 84.6%; Pred. No. 1.3e-32;
 Matches 66; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTPCQEWAAQEPHRHSFTPGTNKWALEKKNYCRNPDGIND 63
 DB 1 CMFGNGKGYRGKATVTGTPCQDWAAQEPHRHSITPTETNPRAGLEKKNYCRNPDGVDG 60

QY 64 PWCYTMNPKLFYCDIP 81
 DB 61 PWCYTMNPKLYDCVP 78

RESULT 4
 AAW07546
 ID AAW07546 standard; protein; 79 AA.
 XX AC AAW07546;
 XX DT 21-JUN-1997 (first entry)
 XX DE Rhesus kringle 1.
 XX KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 XX KW macular degeneration; diabetic retinopathy.
 XX OS Rhesus.
 XX FN W09635774-A2.
 XX PD 14-NOV-1996.
 XX PF 26-APR-1996; 96WO-US05856.
 XX PR 08-MAR-1996; 96US-0612788.
 XX PR 26-APR-1995; 95US-0429743.
 XX PR 22-FEB-1996; 96US-0605598.
 XX PA (CHIL-) CHILDRENS MEDICAL CENT.
 XX PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
 XX DR WPI; 1996-518662/51.
 XX PS Use of angiotatin fragments or aggregates - for inhibiting
 XX PT endothelial cell proliferation and treating angiogenesis-mediated
 XX PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX

Claim 4; Pages 104-105; 203pp; English.

The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotatin fragment, a combination of angiotatin fragments, or
 CC aggregate angiotatin. The fragment is preferably derived from murine,
 CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotatin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, bovine kringle 1, is a specific angiotatin
 CC fragment which can be used in the invention, and represents amino

CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, Rhesus kringle 1, is a specific angiotatin
 CC fragment which can be used in the invention, and represents amino
 CC acids 6-84 of Rhesus angiotatin.
 XX Sequence 79 AA;
 SQ

Query Match 54.0%; Score 274.5; DB 17; Length 79;
 Best Local Similarity 58.8%; Pred. No. 9.4e-20;
 Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTPCQEWAAQEPHRHSFTPGTNKWALEKKNYCRNPDGIND 63
 DB 1 CKTGNGKNRYGNSKTRTGITCQKWSSTSPHR-PTFSATHPSGLEENICRNPDGOG 59

QY 64 PWCYTMNPKLFYCDIFLC 83
 DB 60 PWCYTTDPEERFDYCDIEC 79

RESULT 5
 AAW07548
 ID AAW07548 standard; protein; 79 AA.
 XX AC AAW07548;
 XX DT 21-JUN-1997 (first entry)
 XX DE Bovine kringle 1.
 XX KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 XX KW macular degeneration; diabetic retinopathy.
 XX OS Bos taurus.
 XX XX W09635774-A2.
 XX PD 14-NOV-1996.
 XX PF 26-APR-1996; 96WO-US05856.
 XX PR 08-MAR-1996; 96US-0612788.
 XX PR 26-APR-1995; 95US-0429743.
 XX PR 22-FEB-1996; 96US-0605598.
 XX PA (CHIL-) CHILDRENS MEDICAL CENT.
 XX PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
 XX DR WPI; 1996-518662/51.
 XX PS Use of angiotatin fragments or aggregates - for inhibiting
 XX PT endothelial cell proliferation and treating angiogenesis-mediated
 XX PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX

Claim 4; Page 106; 203pp; English.

The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotatin fragment, a combination of angiotatin fragments, or
 CC aggregate angiotatin. The fragment is preferably derived from murine,
 CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotatin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, bovine kringle 1, is a specific angiotatin
 CC fragment which can be used in the invention, and represents amino

QY 63 GPWCYTMNPKLFYDCDIPLC 83
| | | | | : | : | | | | |
Db 64 GPWCYTTDPKRYDYCDILEC 84

RESULT 8
AAB01906
ID AAB01906 standard; Protein; 84 AA.
AC AAB01906;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen kringle 1 (Tyr80-Glu163).
XX
KW Plasminogen; human; kringle domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antiporiatic; antiinflammatory; antitumor; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS Homo sapiens.
XX
PN US6057122-A.
XX
PD 02-MAY-2000.
XX
PE 05-MAY-1997; 97US-0851350.
XX
PR 03-MAY-1996; 96US-0643219.
XX
PR 03-APR-1997; 97US-0832087.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Davidson DJ;
XX
XX WPI; 2000-349573/30.
XX
XX Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
PS Example 17; Page -: 48pp; English.

CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract,
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC in an exemplification of the invention.
CC Note: This sequence is not shown in the specification, but is derived
CC from the full length human plasminogen sequence (AAB01887) shown in
CC figure 1.
XX
XX Sequence 84 AA;
SQ

Query Match 50.9%; Score 258.5; DB 21; Length 84;
Best Local Similarity 54.3%; Pred. No. 3.8e-18;
Matches 44; Conservative 9; Mismatches 27; Indels 1; Gaps 1;

QY 3 DCMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSSTFIPTGNKWAQLEKNYCRNPDGDN 62
| | | | | : | : | | | | |
Db 4 ECKTGNKNYRGTMSTKNGITCKQWSSTSPHR-PRFSPATHPSEGLEENYCRNPDNDPQ 62
63 GPWCYTMNPKLFYDCDIPLC 83
| | | | | : | : | | | | |
Db 64 GPWCYTTDPKRYDYCDILEC 83

RESULT 9
AAB07545
ID AAB07545 standard; protein; 79 AA.
XX
AC AAB07545;
XX
DT 21-JUN-1997 (first entry)
XX
DE human kringle 1.
XX
KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
KW macular degeneration; diabetic retinopathy.
XX
OS Homo sapiens.
XX
PN WO9635774-A2.
XX
PD 14-NOV-1996.
XX
PE 26-APR-1996; 96WO-US05856.
XX
PR 08-MAR-1996; 96US-0612788.
XX
PR 26-APR-1995; 95US-0429743.
XX
PR 22-FEB-1996; 96US-0605598.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
XX WPI; 1996-518662/51.
XX
PT Use of angiotatin fragments or aggregates - for inhibiting
PT endothelial cell proliferation and treating angiogenesis-mediated
PT diseases, e.g. cancer, arthritis or diabetic retinopathy
XX
PS Claim 4; Page 104; 203pp; English.
XX
CC The invention relates to new methods and compositions for
CC inhibiting endothelial cell proliferation, using as active component
CC an angiotatin fragment, a combination of angiotatin fragments, or
CC aggregate angiotatin. The fragment is preferably derived from murine,
CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
CC 1-4 or kringle 1-4BKL protein. The aggregate angiotatin has a Mol. wt.
CC of 45-65 kD and is derived from a plasminogen fragment beginning at
CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
CC bovine plasminogen. The active component can be used for treating
CC angiogenesis-mediated diseases such as cancer, arthritis, macular
CC degeneration and diabetic retinopathy. It can also be used to develop
CC antibodies for use in diagnosis, detection and therapy.
CC The present sequence, human kringle 1, is a specific angiotatin
CC fragment which can be used in the invention, and represents amino
CC acids 6-84 of human angiotatin.

Query Match 50.5%; Score 256.5; DB 17; Length 79;
Best Local Similarity 55.0%; Pred. No. 5.6e-18;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;
QY 4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSSTFIPTGNKWAQLEKNYCRNPDGDN 63
| | | | | : | : | | | | |
Db 1 CKTGNKNYRGTMSTKNGITCKQWSSTSPHR-PRFSPATHPSEGLEENYCRNPDNDPQ 59


```

QY      64  PWCYTMNPKLFYCDIPLC 83
DB      60  PWCYTTDPKRYDYCDILEC 79

RESULT 10
AAP90598
ID  AAP90598 standard; protein; 81 AA.
XX  AAP90598;
XX
DT  25-MAR-2003 (updated)
DT  31-OCT-2002 (updated)
DT  13-JUN-1990 (first entry)
XX
DE  Synthetic N-terminal protein of truncated human tissue plasminogen
DE  activator (t-PA).
XX
KW  t-PA; tissue plasminogen activator; thrombolytic;
KW  fibrinolytic; fibrin; ds;
XX
OS  Synthetic.
XX
PN  WO8900191-A.
XX
PD  12-JAN-1989.
XX
PF  06-JUL-1988; 88WO-US02263.
XX
PR  06-JUL-1987; 87US-0070012.
PR  20-APR-1988; 88US-0184121.
XX
PA  (GENY ) GENETICS INST INC.
XX
PI  Larsen GR, Ahern T, Langersafe P;
DR  WPT; 1989-039648/05.
DR  N-PSDB; AAN91194.
XX
PT  Tissue plasminogen activator variants -
PT  having N-terminal domain replaced with first kringle region
PT  domain of plasminogen.
XX
PS  Claim 1; Page 32; 36pp; English.
XX
CC  Inserted in front of an N-terminal 1-82 AA truncated t-PA protein, new
CC  product has an increased affinity to fibrin, decreased reactivity with
CC  inhibitors and increased thrombolytic and fibrinolytic activity.
CC  (Updated on 31-OCT-2002 to add missing OS field.)
CC  (Updated on 25-MAR-2003 to correct PA field.)
CC  (Updated on 25-MAR-2003 to correct PI field.)
CC  (Updated on 25-MAR-2003 to correct DR field.)
XX
SQ  Sequence 81 AA;

Query Match      50.0%; Score 254; DB 10; Length 81;
Best Local Similarity 53.1%; Pred. No. 1e-17;
Matches 43; Conservative 9; Mismatches 27; Indels 2; Gaps 1;

QY      3  DCMFNGKGYRGKATVTGTPQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGIN 62
DB      2  ECKTGNKGYRTGSKTKNGATCKWSTSP--HPRFSPATHDSEGLEENYCRNPDNPQ 59

QY      63  GPWCYTMNPKLFYCDIPLC 83
DB      60  GPWCYTTDPKRYDYCDILEC 80

RESULT 11
AAB08410
ID  AAB08410 standard; Protein; 84 AA.
XX

```

```

AC      AAB08410;
XX
DT  20-DEC-2000 (first entry)
XX
DE  Amino acid sequence of kringle 1 of human plasminogen.
XX
KW  Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
KW  tissue plasminogen activator; tumour; atherosclerosis; arthritis;
KW  retinopathy; bronchial vascular congestion; inflammatory bowel disease;
KW  adult respiratory distress syndrome; Castlemans disease; psoriasis;
KW  hepatitis; aneurysm; renal disease; haemangioma.
XX
OS  Homo sapiens.
XX
PN  WO2000049871-A1.
XX
PD  31-AUG-2000.
XX
PF  24-FEB-2000; 2000WO-US04798.
XX
PR  24-FEB-1999; 99US-0121341.
PR  25-FEB-1999; 99US-0121633.
PR  18-NOV-1999; 99US-0166176.
XX
PA  (FORD-) FORD HEALTH SYSTEM HENRY.
XX
PI  Dou D, Chopp M, Wang L, Mikkelsen T;
XX
PF  WPI; 2000-572016/53.
XX
PT  Use of kringle protein and kringle derived from plasminogen and
PT  composition comprising kringle proteins for treating tumor and
PT  atherosclerosis, arthritis and retinopathy -
XX
PS  Disclosure; Fig 6; 163pp; English.
XX
CC  The specification describes a human polypeptide which is a potent
CC  angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle
CC  proteins, or a kringle derived from human tissue plasminogen activator
CC  (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
CC  or tPA are useful for treating tumours, as well as atherosclerosis,
CC  arthritis, retinopathy and other similar diseases. KED is also useful
CC  for the treatment of diseases such as bronchial vascular congestion,
CC  inflammatory bowel disease, psoriasis, hepatitis, aneurysm, renal disease
CC  and haemangioma. The present sequence represents kringle 1 of human
CC  plasminogen, which is used in the course of the invention.
XX
SQ  Sequence 84 AA;

Query Match      49.7%; Score 252.5; DB 21; Length 84;
Best Local Similarity 53.1%; Pred. No. 1.5e-17;
Matches 43; Conservative 8; Mismatches 29; Indels 1; Gaps 1;

QY      3  DCMFNGKGYRGKATVTGTPQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGIN 62
DB      3  ECKTGNKGYRTGSKTKNGATCKWSTSPHR--PRFSPATHDSEGLEENYCRNPDNPQ 61

QY      63  GPWCYTMNPKLFYCDIPLC 83
DB      62  GPWCYTTDPKRYDYCDILEC 82

RESULT 12
AAP81360
ID  AAP81360 standard; protein; 79 AA.
XX
AC  AAP81360;
XX
DT  25-MAR-2003 (updated)
DT  31-OCT-2002 (updated)
DT  06-OCT-1990 (first entry)
XX

```

```

DE  KI domain of plasminogen.
XX
XX  Tissue plasminogen activator; analogue; K1 domain.
XX
XX  Homo sapiens.
XX
XX  EF293934-A.
XX
XX  07-DEC-1988.
XX
XX  03-JUN-1988; 88EP-0108949.
XX
XX  04-JUN-1987; 87US-0058217.
XX
XX  (ZYMO ) ZYMOGENETICS INC.
XX  (NOVO ) NOVO IND AS.
XX  (EISA ) EISA CO LTD.
XX
XX  Mulvihill ER, Nexo BA, Yoshitake S, Ikeda Y, Suzuki S;
XX  Hashimoto A, Yuzuriha T;
XX
XX  WPI; 1988-347625/49.
XX  N-PSDB; AAN81087.
XX
XX  New tissue plasminogen activator analogues with non-native K1-domain -
XX  having higher fibrin specificity, and new encoding DNA sequences.
XX
XX  Disclosure; Page ?; ?pp; English.
XX
XX  The K1 domain of plasminogen may be used to replace the K1 domain of
XX  t-PA to produce an analogue with greater specificity for fibrin than
XX  native t-PA.
XX
XX  (Updated on 31-OCT-2002 to add missing OS field.)
XX  (Updated on 25-MAR-2003 to correct PA field.)
XX  (Updated on 25-MAR-2003 to correct PI field.)
XX
XX  Sequence 79 AA;

Query Match          49.5%; Score 251.5; DB 9; Length 79;
Best Local Similarity 53.8%; Pred. No. 1.7e-17;
Matches 43; Conservative 9; Mismatches 27; Indels 1; Gaps 1;

QY  4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNCRNPDGDI 63
DB  1 CKTGDKNYRGTMSKTKNGITCKQWSSTSFPHR-PRFSPATHPSEGLEENYCRNPDNDPQG 59

QY  64 PWCYTMNPKLFYCDIPLC 83
DB  60 PWCYTTDPKRYDYCDILEC 79

RESULT 13
AAR96221
ID  AAR96221 standard; Protein; 79 AA.
XX
XX  AAR96221;
XX
XX  25-MAR-2003 (updated)
XX  06-AUG-1996 (first entry)
XX
XX  Plasminogen kringle domain K1.
XX
XX  Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
XX  fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
XX  protein engineering; kringle.
XX
XX  Synthetic.
XX
XX  US5504001-A.
XX
XX  02-APR-1996.
XX
XX  06-JUN-1994; 94US-0254485.

DE  K1 domain (AAR96221) of plasminogen was incorporated
XX  into tissue plasminogen activator (tPA) (see also AAR96220) as a
XX  replacement for the native tPA kringle 1 sequence, producing
XX  a novel plasminogen activator (see also AAR96222). This was
XX  performed by insertion of plasminogen K1 DNA (AAT27586) into
XX  a native tPA sequence (AAT27585) and expression in Escherichia
XX  coli RRI transformants (FERM P-9272). Such protocols can be used
XX  to produce novel plasminogen activators that show increased clot
XX  lysing specificity or plasma half-life.
XX  (Updated on 25-MAR-2003 to correct PF field.)
XX
XX  Sequence 79 AA;

Query Match          49.5%; Score 251.5; DB 17; Length 79;
Best Local Similarity 53.8%; Pred. No. 1.7e-17;
Matches 43; Conservative 9; Mismatches 27; Indels 1; Gaps 1;

QY  4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNCRNPDGDI 63
DB  1 CKTGDKNYRGTMSKTKNGITCKQWSSTSFPHR-PRFSPATHPSEGLEENYCRNPDNDPQG 59

QY  64 PWCYTMNPKLFYCDIPLC 83
DB  60 PWCYTTDPKRYDYCDILEC 79

RESULT 14
AAR94401
ID  AAR94401 standard; Protein; 78 AA.
XX
XX  AAR94401;
XX
XX  01-JUL-2002 (updated)
XX  18-JUN-1990 (first entry)
XX
XX  Sequence encoding K1 domain of plasminogen.
XX
XX  Tissue plasminogen activator; fibrinolytic; t-PA; plasminogen.
XX  Homo sapiens.
XX
XX  AU8817430-A.
XX
XX  08-DEC-1988.
XX
XX  06-JUN-1988; 88AU-0017430.
XX
XX  04-JUN-1987; 87JP-0058061.
XX  02-MAR-1988; 88JP-0062847.
XX
XX  (NOVO ) NOVO INDUSTI A/S.
XX  (ZYMO ) ZYMOGENETICS INC.
XX  (EISA ) EISA CO LTD.
XX
XX  Mulvihill ER, Ohara PJ, Nexo BA, Yoshitake S, Ikeda Y, Suzuki S;

```

PI Hashimoto A, Yuzuhira T;
 XX WPI; 1989-039773/06.
 DR N-PSDB; AAN91218.
 XX
 PT Tissue plasminogen activator analogues -
 PT having at least one cysteine residue in the growth factor domain
 PT replaced with another amino acid.
 XX
 PS Example 9; Fig 11; 95pp; English.
 XX
 CC Amino acid sequence of kringle domain of plasminogen, incorporated into
 CC plasmid PK1. Sequence is derived from 11 synthetic oligonucleotides.
 CC Alternatively the sequence is described as encoding Asn at position 96.
 CC (Updated on 01-JUL-2002 to add missing PI field.)
 XX
 SQ Sequence 78 AA;
 Query Match 48.9%; Score 248.5; DB 10; Length 78;
 Best Local Similarity 54.5%; Pred. No. 3.4e-17;
 Matches 42; Conservative 9; Mismatches 25; Indels 1; Gaps 1;
 QY 4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNCYCRNPDGDDING 63
 Db 1 CKTGDKNGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNCYCRNPDGDDING 63
 QY 64 PWCYTMDPKLFYDCDI 80
 Db 60 PWCYTMDPKLFYDCDI 76
 RESULT 15
 AAW07547
 ID AAW07547 standard; protein; 79 AA.
 AC AAW07547;
 XX
 XX 21-JUN-1997 (first entry)
 XX
 DE Porcine kringle 1.
 XX
 KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 KW macular degeneration; diabetic retinopathy.
 XX
 OS Sus scrofa.
 XX
 XX WO9635774-A2.
 XX
 PD 14-NOV-1996.
 XX
 XX 26-APR-1996; 96WO-0505856.
 XX
 XX 08-MAR-1996; 96US-0612788.
 XX
 XX 26-APR-1995; 95US-0429743.
 XX
 XX 22-FEB-1996; 96US-0605598.
 XX
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 XX
 XX Cao Y, Folkman MJ, Lin J, Oreilly MS, Sim KL;
 XX WPI; 1996-518662/51.
 XX
 PT Use of angiotatin fragments or aggregates - for inhibiting
 PT endothelial cell proliferation and treating angiogenesis-mediated
 PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX
 XX Claim 4; Page 105; 203pp; English.
 XX
 CC The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotatin fragment, a combination of angiotatin fragments, or
 CC aggregate angiotatin. The fragment is preferably derived from murine,
 CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,

CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotatin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, porcine kringle 1, is a specific angiotatin
 CC fragment which can be used in the invention, and represents amino
 CC acids 6-84 of porcine angiotatin.
 XX
 SQ Sequence 79 AA;

Query Match 48.5%; Score 246.5; DB 17; Length 79;
 Best Local Similarity 53.8%; Pred. No. 5.4e-17;
 Matches 43; Conservative 7; Mismatches 29; Indels 1; Gaps 1;
 QY 4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNCYCRNPDGDDING 63
 Db 1 CKTGNGKNGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNCYCRNPDGDDING 59
 QY 64 PWCYTMDPKLFYDCDI 83
 Db 60 PWCYTMDPKLFYDCDI 79

Search completed: September 29, 2003, 08:46:58
 Job time : 83 secs


```
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/09/066,028
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/612,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: K1
/
/ US-09-066-028-9
/
/ Query Match 54.0%; Score 274.5; DB 3; Length 79;
/ Best Local Similarity 58.8%; Pred. NO. 3.9e-24;
/ Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;
/
/ QY 4 CMFGNGKGYRGKKATVTGTPCOEWAAQEPHRHSTFIPGINKWAGLEKKNYCRNPDGDLG 63
/ Db 1 CKTGNGKNYRGTKMTGTGTCOKWSSYSPHR-PTFSPATHPSEGLEENYCRNPDNDGQG 59
/
/ QY 64 PWCYTMNPKLFYDCDIPLC 83
/ Db 60 PWCYTTDPERFYCDIPEC 79
/
/ RESULT 5
/ US-09-335-325-9
/ Sequence 9, Application US/09335325
/ Patent No. 6521439
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ O'Reilly, Micheal
/ Cao, Yihai
/ Sim, B. Kim Lee
/
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S
/ ZIP: 30303-1769
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
```

```
/
/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: K1
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 9:
/
/ US-09-335-325-9
/
/ Query Match 54.0%; Score 274.5; DB 4; Length 79;
/ Best Local Similarity 58.8%; Pred. NO. 3.9e-24;
/ Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;
/
/ QY 4 CMFGNGKGYRGKKATVTGTPCOEWAAQEPHRHSTFIPGINKWAGLEKKNYCRNPDGDLG 63
/ Db 1 CKTGNGKNYRGTKMTGTGTCOKWSSYSPHR-PTFSPATHPSEGLEENYCRNPDNDGQG 59
/
/ QY 64 PWCYTMNPKLFYDCDIPLC 83
/ Db 60 PWCYTTDPERFYCDIPEC 79
/
/ RESULT 6
/ US-08-612-788-11
/ Sequence 11, Application US/08612788
/ Patent No. 5837682
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ O'Reilly, Micheal
/ Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S
/ ZIP: 30303-1769
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE:
```

```
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Bovine
/ IMMEDIATE SOURCE:
/ CLONE: K1
/
/ US-08-612-788-11
/
/ Query Match 50.9%; Score 258.5; DB 2; Length 79;
/ Best Local Similarity 57.5%; Pred. No. 2.6e-22;
/ Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;
/
/ QY 4 CMFGNGKGYRGKATVTGTCQEWAAQPHRHSFIPTGKNWAGLEKNCVCRNPDDG 63
/ DB 1 CKTGNGQTYRGTAETKSGVTCQKWSATSPH-VPKFSPEKFPFLAGLEENYCRNPNDENG 59
/
/ QY 64 PWCYTMNPKLFYDCIPLC 83
/ DB 60 PWCYTTDPKRYDYCDIPEC 79
/
/ RESULT 7
/ US-09-066-028-11
/ Sequence 11, Application US/09066028
/ Patent No. 6024688
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/066,028
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/612,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
```

```
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Bovine
/ IMMEDIATE SOURCE:
/ CLONE: K1
/
/ US-09-066-028-11
/
/ Query Match 50.9%; Score 258.5; DB 3; Length 79;
/ Best Local Similarity 57.5%; Pred. No. 2.6e-22;
/ Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;
/
/ QY 4 CMFGNGKGYRGKATVTGTCQEWAAQPHRHSFIPTGKNWAGLEKNCVCRNPDDG 63
/ DB 1 CKTGNGQTYRGTAETKSGVTCQKWSATSPH-VPKFSPEKFPFLAGLEENYCRNPNDENG 59
/
/ QY 64 PWCYTMNPKLFYDCIPLC 83
/ DB 60 PWCYTTDPKRYDYCDIPEC 79
/
/ RESULT 8
/ US-09-335-325-11
/ Sequence 11, Application US/09335325
/ Patent No. 6521439
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
```

```
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Bovine
/ IMMEDIATE SOURCE:
/ CLONE: K1
/ SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-335-325-11

Query Match          50.9%; Score 258.5; DB 4; Length 79;
Best Local Similarity 57.5%; Pred. No. 2.6e-22;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY 4 CMFNGKGYRGKATVTTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGDI 63
Db 1 CKTGNGQYRGTTAETKSGVTCOKWSATSPH-VPKTSPEKFLAGLEENYCRNPDND 59
QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCITTDPRYDYCDILEC 79

RESULT 9
US-09-348-953-1
/ Sequence 1, Application US/09348953
/ Patent No. 6538103
/ GENERAL INFORMATION:
/ APPLICANT: Ji, Richard Wei-Dong
/ APPLICANT: Trail, Pamela A.
/ TITLE OF INVENTION: LYSINE BINDING FRAGMENTS OF ANGIOSTATIN
/ FILE REFERENCE: DB11sequences
/ CURRENT APPLICATION NUMBER: US/09/348,953
/ CURRENT FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: 60/092,831
/ PRIOR FILING DATE: 1998-07-14
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 82
/ TYPE: PRT
/ ORGANISM: human
US-09-348-953-1

Query Match          50.9%; Score 258.5; DB 4; Length 82;
Best Local Similarity 54.3%; Pred. No. 2.7e-22;
Matches 44; Conservative 9; Mismatches 27; Indels 1; Gaps 1;

QY 3 DCMFNGKGYRGKATVTTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGDI 62
Db 2 CKTGNGKNGYRGTMSTKNGITCQKWSSTSPHR-PFSPATHPSEGLEENYCRNPDND 60
QY 63 GPWCYTMNPKLFYDCDIPLC 83
Db 61 GPWCYTTDPKRYDYCDILEC 81

RESULT 10
US-08-612-788-8
/ Sequence 8, Application US/08612788
/ Patent No. 5837682
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Michael
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
```

```
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612.788
/ FILING DATE:
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K1
US-08-612-788-8

Query Match          50.5%; Score 256.5; DB 2; Length 79;
Best Local Similarity 55.0%; Pred. No. 4.4e-22;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFNGKGYRGKATVTTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKNGYRGTMSTKNGITCQKWSSTSPHR-PFSPATHPSEGLEENYCRNPDND 59
QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCITTDPRYDYCDILEC 79

RESULT 11
US-08-763-528A-2
/ Sequence 2, Application US/08763528A
/ Patent No. 5854221
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yihai
/ APPLICANT: Folkman, M. Judah
/ TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
/ TITLE OF INVENTION: and Method of Use
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew, LLP
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: US
/ ZIP: 30303
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/763,528A
/ FILING DATE: 12-DEC-1996
```


; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K1
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Query Match 50.5%; Score 256.5; DB 4; Length 79;
Best Local Similarity 55.0%; Pred. No. 4.4e-22;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;
QY 4 CMFNGKGYRGKATTTGTGTCQEWAAQEPHRRHSTFIPGTNKGAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKRYGTMTSKNGITCKQWSSTSPHR-PREFSPATHPSEGLEENYCRNPDNDPQG 59
QY 64 PWCYTMNPKLFYDCDIPIC 83
Db 60 PWCYTTDPETRFDYCDIPEC 79

RESULT 14
US-08-612-788-10
; Sequence 10, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Porcine
; IMMEDIATE SOURCE:
; CLONE: K1
US-08-612-788-10

Query Match 48.5%; Score 246.5; DB 2; Length 79;
Best Local Similarity 53.8%; Pred. No. 6e-21;
Matches 43; Conservative 7; Mismatches 29; Indels 1; Gaps 1;
QY 4 CMFNGKGYRGKATTTGTGTCQEWAAQEPHRRHSTFIPGTNKGAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKRYGTMTSKNGITCKQWSSTSPH-IPKYSPEKFLAGLEENYCRNPDNDKG 59
QY 64 PWCYTMNPKLFYDCDIPIC 83
Db 60 PWCYTTDPETRFDYCDIPEC 79

RESULT 15
US-09-066-028-10
; Sequence 10, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Porcine
; IMMEDIATE SOURCE:
; CLONE: K1
US-09-066-028-10

Query Match 48.5%; Score 246.5; DB 3; Length 79;
Best Local Similarity 53.8%; Pred. No. 6e-21;
Matches 43; Conservative 7; Mismatches 29; Indels 1; Gaps 1;
QY 4 CMFNGKGYRGKATTTGTGTCQEWAAQEPHRRHSTFIPGTNKGAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKRYGTMTSKNGITCKQWSSTSPH-IPKYSPEKFLAGLEENYCRNPDNDKG 59

Db 1 CKTGNGKNYRGTTSTKXSGVICQKWSVSPH-IPKISPEKFPPLAGLEENYCRNPNDDEKG 59

Qy 64 PWCYTNNPKLFDYCDIPLC 83

||||| : | ||||| |

Db 60 PWCYTTPETRFDYCDIPEC 79

Search completed: September 29, 2003, 08:50:31
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:49:20 ; Search time 63 Seconds

(without alignments)
206.546 Million cell updates/sec

Title: US-10-088-548-8

Perfect score: 508

Sequence: 1 BQDCMFGNGYGRKATTV.....YTNPKRLFDYCDIPLCASS 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 198464

Minimum DB seq length: 0

Maximum DB seq length: 86

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

```
1: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pdb.pdb.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pdb.pdb.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pdb.pdb.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pdb.pdb.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pdb.pdb.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pdb.pdb.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pdb.pdb.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pdb.pdb.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pdb.pdb.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pdb.pdb.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pdb.pdb.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pdb.pdb.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pdb.pdb.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pdb.pdb.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pdb.pdb.*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pdb.pdb.*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pdb.pdb.*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pdb.pdb.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412	81.1	80	9	US-09-753-064-6
2	412	81.1	80	12	US-10-267-137-8
3	405	79.7	79	9	US-09-753-064-1
4	387	76.2	80	9	US-09-761-120-47
5	274.5	54.0	79	9	US-09-761-120-9
6	274.5	54.0	79	10	US-09-335-325-9
7	274.5	54.0	79	15	US-10-131-241-9
8	258.5	50.9	79	9	US-09-761-120-11
9	258.5	50.9	79	10	US-09-335-325-11
10	258.5	50.9	79	15	US-10-131-241-11
11	256.5	50.5	79	9	US-09-753-064-2
12	256.5	50.5	79	9	US-09-761-120-8
13	256.5	50.5	79	10	US-09-335-325-8
14	256.5	50.5	79	12	US-10-267-137-7
15	256.5	50.5	79	15	US-10-131-241-8

```
16 246.5 48.5 79 9 US-09-761-120-10
17 246.5 48.5 79 10 US-09-335-325-10
18 246.5 48.5 79 15 US-10-131-241-10
19 241.5 47.5 79 10 US-09-335-325-7
20 241.5 47.5 79 15 US-10-131-241-7
21 240 47.2 78 9 US-09-761-120-22
22 240 47.2 78 10 US-09-335-325-22
23 240 47.2 78 15 US-10-131-241-22
24 232.5 45.8 79 9 US-09-761-120-7
25 231.5 45.6 79 12 US-10-267-137-10
26 231.5 45.6 80 12 US-10-267-137-1
27 222 43.7 78 9 US-09-753-064-5
28 222 43.7 78 9 US-09-761-120-17
29 222 43.7 78 9 US-09-761-120-23
30 222 43.7 78 10 US-09-335-325-17
31 222 43.7 78 10 US-09-335-325-23
32 222 43.7 78 12 US-10-267-137-9
33 222 43.7 78 15 US-10-131-241-17
34 222 43.7 78 15 US-10-131-241-23
35 220 43.3 78 9 US-09-761-120-16
36 220 43.3 78 9 US-09-761-120-20
37 220 43.3 78 10 US-09-335-325-16
38 220 43.3 78 10 US-09-335-325-20
39 220 43.3 78 15 US-10-131-241-16
40 220 43.3 78 15 US-10-131-241-20
41 217 42.7 78 9 US-09-761-120-12
42 217 42.7 78 10 US-09-335-325-12
43 217 42.7 78 15 US-10-131-241-12
44 216 42.5 78 9 US-09-761-120-14
45 216 42.5 78 10 US-09-335-325-13
```

ALIGNMENTS

```
RESULT 1
US-09-753-064-6
; Sequence 6, Application US/09753064
; Patent No. US20010016644A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; Folkmann, M. Judah
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
; and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,064
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 12-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
```

```
Sequence 10, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 7, Appl
Sequence 10, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 17, Appl
Sequence 23, Appl
Sequence 17, Appl
Sequence 23, Appl
Sequence 17, Appl
Sequence 9, Appl
Sequence 17, Appl
Sequence 23, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 16, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 14, Appl
Sequence 13, Appl
```

```
;
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..80
; OTHER INFORMATION: /note= "Kringle 5 - Figure 3"
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-753-064-6

Query Match      81.1%; Score 412; DB 9; Length 80;
Best Local Similarity 83.8%; Pred. No. 6.5e-39;
Matches 67; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFPGTNKAGLEKNCRNPDGDDING 63
Db 1 CMFGNGKGYRGKATVTGTPCQDWAAQEPHRSHTFPGTNPRAGLEKNCRNPDGDDVGG 60

QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 61 PWCYTMNPKLYDCDVPQC 80

RESULT 2
US-10-267-137-8
; Sequence 8, Application US/10267137
; Publication No. US20030148950A1
; GENERAL INFORMATION:
; APPLICANT: Xin, Li
; APPLICANT: Li, Zai-Ping
; APPLICANT: Gan, Ren-bao
; APPLICANT: Zhou, Qing-wei
; APPLICANT: Xu, Ren
; TITLE OF INVENTION: KRINGLE DOMAIN 1 OF HUMAN HEPATOCYTE
; FILE REFERENCE: 524282000100
; CURRENT APPLICATION NUMBER: US/10/267,137
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/328,329
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 80
; TYPE: PRI
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence alignment of kringle domains of
; OTHER INFORMATION: plasminogen and HGF
US-10-267-137-8

Query Match      81.1%; Score 412; DB 12; Length 80;
Best Local Similarity 83.8%; Pred. No. 6.5e-39;
Matches 67; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFPGTNKAGLEKNCRNPDGDDING 63
Db 1 CMFGNGKGYRGKATVTGTPCQDWAAQEPHRSHTFPGTNPRAGLEKNCRNPDGDDVGG 60

QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 61 PWCYTMNPKLYDCDVPQC 80

RESULT 3
US-09-753-064-1
; Sequence 1, Application US/09753064
; Patent No. US2001001664A1
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; FOLKMAN, M. Judah
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
; and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,064
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/763,528
; FILING DATE: 12-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-753-064-1

Query Match      79.7%; Score 405; DB 9; Length 79;
Best Local Similarity 84.6%; Pred. No. 3.9e-38;
Matches 66; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFPGTNKAGLEKNCRNPDGDDING 63
Db 1 CMFGNGKGYRGKATVTGTPCQDWAAQEPHRSHTFPGTNPRAGLEKNCRNPDGDDVGG 60

QY 64 PWCYTMNPKLFYDCDIP 81
Db 61 PWCYTMNPKLYDCDVP 78

RESULT 4
US-09-761-120-47
; Sequence 47, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
```

```
/ NUMBER OF SEQ ID NOS: 47
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 47
/ LENGTH: 80
/ TYPE: PRT
/ ORGANISM: Murinae gen. sp.
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Kringle 5
US-09-761-120-47

Query Match          76.2%; Score 387; DB 9; Length 80;
Best Local Similarity 80.0%; Pred. No. 4.2e-36;
Matches 64; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTPCOEWAAQEPHRSHTFIPGINKWAGLEKNCYCRNPDGNG 63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CMYGNKDYRGKATVAAGTGPCQGWAAQEPHRSHTFPTQNPRADEKNCYCRNPDGNG 60

QY 64 PWCYTNPKLFEDYCDIPLC 83
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 PWCYTNPKLFEDYCDIPLC 80

RESULT 5
US-09-761-120-9
/ Sequence 9, Application US/09761120
/ Patent No. US20020037847A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Michael
/ TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
/ FILE REFERENCE: 05940-0151 (43171-252068)
/ CURRENT APPLICATION NUMBER: US/09/761,120
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: 09/309,821
/ PRIOR FILING DATE: 1999-05-11
/ PRIOR APPLICATION NUMBER: 08/866,735
/ PRIOR FILING DATE: 1997-05-30
/ NUMBER OF SEQ ID NOS: 47
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 9
/ LENGTH: 79
/ TYPE: PRT
/ ORGANISM: Macaca sp.
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Kringle 1
US-09-761-120-9

Query Match          54.0%; Score 274.5; DB 9; Length 79;
Best Local Similarity 58.8%; Pred. No. 1.7e-23;
Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTPCOEWAAQEPHRSHTFIPGINKWAGLEKNCYCRNPDGNG 63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CKTGNGKNGYRGKATVTGTPCOEWAAQEPHRSHTFIPGINKWAGLEKNCYCRNPDGNG 59

QY 64 PWCYTNPKLFEDYCDIPLC 83
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 60 PWCYTNPKLFEDYCDIPLC 80

RESULT 6
US-09-335-325-9
/ Sequence 9, Application US/09335325
/ Patent No. US20020164717A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Michael
/ APPLICANT: Cao, Fihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
```

```
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: KI
/ SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-335-325-9

Query Match          54.0%; Score 274.5; DB 10; Length 79;
Best Local Similarity 58.8%; Pred. No. 1.7e-23;
Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTPCOEWAAQEPHRSHTFIPGINKWAGLEKNCYCRNPDGNG 63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CKTGNGKNGYRGKATVTGTPCOEWAAQEPHRSHTFIPGINKWAGLEKNCYCRNPDGNG 59

QY 64 PWCYTNPKLFEDYCDIPLC 83
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 60 PWCYTNPKLFEDYCDIPLC 79

RESULT 7
US-10-131-241-9
/ Sequence 9, Application US/10131241
/ Publication No. US20030012792A1
/ GENERAL INFORMATION:
/ APPLICANT: Holaday, John W.
/ APPLICANT: Fortier, Anne H.
/ TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
/ FILE REFERENCE: 05213-0344 43170-271565
/ CURRENT APPLICATION NUMBER: US/10/131,241
/ CURRENT FILING DATE: 2002-07-22
/ PRIOR APPLICATION NUMBER: US 09/413,049
/ PRIOR FILING DATE: 1999-10-06
/ PRIOR APPLICATION NUMBER: US 09/316,802
/ PRIOR FILING DATE: 1999-05-21
```

```
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Rhesus monkey
US-10-131-241-9

Query Match          54.0%; Score 274.5; DB 15; Length 79;
Best Local Similarity 58.8%; Pred. No. 1.7e-23;
Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY      4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNYCRNPDGDING 63
      1 CKTGNGKNTVGRGWSKTRIGTCQKWSATSPH-PTFSPATHSEGLEENYCRNPDNDGOG 59

Db      64 PWCYTMNPKLFYCDIPLC 83
      60 PWCYTTDPKRYDYCDIPEC 79

RESULT 8
US-09-761-120-11
; Sequence 11, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Bos sp.
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Kringle 1
US-09-761-120-11

Query Match          50.9%; Score 258.5; DB 9; Length 79;
Best Local Similarity 57.5%; Pred. No. 1.1e-21;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY      4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNYCRNPDGDING 63
      1 CKTGNGQTVRGTAETKSGVTCQKWSATSPH-VPKFSPKFFLAGLEENYCRNPDNDENG 59

Db      64 PWCYTMNPKLFYCDIPLC 83
      60 PWCYTTDPKRYDYCDIPEC 79

RESULT 9
US-09-335-325-11
; Sequence 11, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; IMMEDIATE SOURCE:
; CLONE: K1
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-335-325-11

Query Match          50.9%; Score 258.5; DB 10; Length 79;
Best Local Similarity 57.5%; Pred. No. 1.1e-21;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY      4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNYCRNPDGDING 63
      1 CKTGNGQTVRGTAETKSGVTCQKWSATSPH-VPKFSPKFFLAGLEENYCRNPDNDENG 59

Db      64 PWCYTMNPKLFYCDIPLC 83
      60 PWCYTTDPKRYDYCDIPEC 79

RESULT 10
US-10-131-241-11
; Sequence 11, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolif
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
```

```
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Bovine
US-10-131-241-11

Query Match          50.9%; Score 258.5; DB 15; Length 79;
Best Local Similarity 57.5%; Pred. No. 1.1e-21;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATTTGTPCOEWAAQEPHRHSTFIPGTNKWAGLEKNCYCRNPDDING 63
Db 1 CKTGNGQYRGTTATSGTCKNSATSPH-VPFSPKFPPLAGLENYCRNPDDENG 59

QY 64 PWCYTNNPKLFYDCDIPLC 83
Db 60 PWCYTTPDEKRYDCDIPEC 79

RESULT 11
US-09-753-064-2
; Sequence 2, Application US/09753064
; Patent No. US20010016644A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,064
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/763,528
; FILING DATE: 12-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..79
; OTHER INFORMATION: /note= "Kringle 1 - Figure 3"
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
```

```
US-09-753-064-2

Query Match          50.5%; Score 256.5; DB 9; Length 79;
Best Local Similarity 55.0%; Pred. No. 1.8e-21;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATTTGTPCOEWAAQEPHRHSTFIPGTNKWAGLEKNCYCRNPDDING 63
Db 1 CKTGNGKNYRGTMSTKNGITCKWSSTSPHR-PREFSPATHPSGLEENYCRNPDDPQG 59

QY 64 PWCYTNNPKLFYDCDIPLC 83
Db 60 PWCYTTPDEKRYDCDILEC 79

RESULT 12
US-09-761-120-8
; Sequence 8, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmid
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Kringle 1
US-09-761-120-8

Query Match          50.5%; Score 256.5; DB 9; Length 79;
Best Local Similarity 55.0%; Pred. No. 1.8e-21;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATTTGTPCOEWAAQEPHRHSTFIPGTNKWAGLEKNCYCRNPDDING 63
Db 1 CKTGNGKNYRGTMSTKNGITCKWSSTSPHR-PREFSPATHPSGLEENYCRNPDDPQG 59

QY 64 PWCYTNNPKLFYDCDIPLC 83
Db 60 PWCYTTPDEKRYDCDILEC 79

RESULT 13
US-09-335-325-8
; Sequence 8, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
```



```
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K1
/ SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-335-325-8

Query Match 50.5%; Score 256.5; DB 10; Length 79;
Best Local Similarity 55.0%; Pred. No. 1.8e-21;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTPCQEWAAOEPHRHSTFIPGTNKGWAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKNYRGWMTSKNGITCQKWSSTSPHR-PRFSPATHPSEGLEENYCRNPDNDPQ 59

QY 64 PWCYTMNPKLFYCDIPLC 83
Db 60 PWCYTTDPKRYDYCDILEC 79

RESULT 14
US-10-267-137-7
/ Sequence 7, Application US/10267137
/ Publication No. US20030148950A1
/ GENERAL INFORMATION:
/ APPLICANT: Xin, Li
/ APPLICANT: Li, Zai-Ping
/ APPLICANT: Gan, Ren-Bao
/ APPLICANT: Zhou, Qing-wei
/ APPLICANT: Xu, Ren
/ TITLE OF INVENTION: KRINGLE DOMAIN 1 OF HUMAN HEPATOCYTE
/ FILE REFERENCE: 524282000100
/ CURRENT APPLICATION NUMBER: US/10/267,137
/ CURRENT FILING DATE: 2002-10-07
/ PRIOR APPLICATION NUMBER: 60/328,329
/ PRIOR FILING DATE: 2001-10-09
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 79
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Sequence alignment of kringle domains of

/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K1
/ SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-335-325-8

Query Match 50.5%; Score 256.5; DB 10; Length 79;
Best Local Similarity 55.0%; Pred. No. 1.8e-21;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTPCQEWAAOEPHRHSTFIPGTNKGWAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKNYRGWMTSKNGITCQKWSSTSPHR-PRFSPATHPSEGLEENYCRNPDNDPQ 59

QY 64 PWCYTMNPKLFYCDIPLC 83
Db 60 PWCYTTDPKRYDYCDILEC 79

RESULT 15
US-10-131-241-8
/ Sequence 8, Application US/10131241
/ Publication No. US20030012792A1
/ GENERAL INFORMATION:
/ APPLICANT: Holaday, John W.
/ APPLICANT: Fortier, Anne H.
/ TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolif
/ FILE REFERENCE: 05213-0344 43170-271565
/ CURRENT APPLICATION NUMBER: US/10/131,241
/ CURRENT FILING DATE: 2002-07-22
/ PRIOR APPLICATION NUMBER: US 09/413,049
/ PRIOR FILING DATE: 1999-10-06
/ PRIOR APPLICATION NUMBER: US 09/316,802
/ PRIOR FILING DATE: 1999-05-21
/ PRIOR APPLICATION NUMBER: US 60/086,586
/ PRIOR FILING DATE: 1998-05-22
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 8
/ LENGTH: 79
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-131-241-8

Query Match 50.5%; Score 256.5; DB 15; Length 79;
Best Local Similarity 55.0%; Pred. No. 1.8e-21;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTPCQEWAAOEPHRHSTFIPGTNKGWAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKNYRGWMTSKNGITCQKWSSTSPHR-PRFSPATHPSEGLEENYCRNPDNDPQ 59

QY 64 PWCYTMNPKLFYCDIPLC 83
Db 60 PWCYTTDPKRYDYCDILEC 79

Search completed: September 29, 2003, 08:59:02
Job time : 64 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: September 29, 2003, 08:44:24 ; Search time 39 Seconds
(without alignments)
212.064 Million cell updates/sec

Title: US-10-088-548-8
Perfect score: 508
Sequence: 1 EQDCMFGNGKGYRGKATV.....YTMNPKLFDYCDIPLCASS 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168692 residues
Total number of hits satisfying chosen parameters: 29185

Minimum DB seq length: 0
Maximum DB seq length: 86

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:**
1: pir1:**
2: pir2:**
3: pir3:**
4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result. No.	Score	Query Match	Length	ID	Description
1	54.5	10.7	64	1 NTSR5L	neurotoxin V - Egy
2	54.5	10.7	65	2 E82078	conserved hypothet
3	52.5	10.3	64	1 NTSR5M	neurotoxin V - sco
4	51	10.0	65	2 T03082	hypothetical prote
5	49.5	9.7	68	2 G82600	hypothetical prote
6	48.5	9.5	65	1 NTSREB	neurotoxin XI - sc
7	48.5	9.5	65	2 PQ085B	glycoprotein H - h
8	48	9.4	40	2 T35510	hypothetical prote
9	47.5	9.4	40	2 S21332	hypothetical prote
10	47.5	9.4	59	2 D81186	hypothetical prote
11	47	9.3	59	1 T6P5UJ	mambin - eastern J
12	47	9.3	77	2 AF2564	hypothetical prote
13	46.5	9.2	51	2 AF3350	hypothetical prote
14	46.5	9.2	62	2 A69034	hypothetical prote
15	46.5	9.2	73	2 F83824	hypothetical prote
16	46.5	9.2	86	2 T17313	hypothetical prote
17	46	9.1	71	2 A34768	ORF1 protein - Ori
18	45.5	9.0	65	2 PQ0850	glycoprotein H - h
19	45.5	9.0	65	2 PQ0848	hypothetical prote
20	45.5	9.0	65	2 AE2406	telomeric DNA bind
21	45	8.9	46	2 S67923	hypothetical prote
22	45	8.9	80	2 E97800	hypothetical prote
23	44.5	8.8	50	2 E82529	hypothetical prote
24	44.5	8.8	62	2 T02631	hypothetical prote
25	44	8.7	62	2 G64561	hypothetical prote
26	44	8.7	64	2 A35940	neurotoxin alpha-I
27	44	8.7	69	2 T14285	embryogenic callus
28	44	8.7	74	2 A53356	toxin CSRX-1 - wan
29	44	8.7	85	2 A39306	alpha insect toxin

30	44	8.7	86	2 F82496	hypothetical prote
31	43.5	8.6	49	2 S70093	hypothetical prote
32	43.5	8.6	63	2 AE2350	hypothetical prote
33	43.5	8.6	70	2 AC0718	probable cold shoc
34	43.5	8.6	86	2 G69452	molybdopterin conv
35	43	8.5	52	2 T32835	hypothetical prote
36	43	8.5	59	2 S39068	homeotic protein c
37	43	8.5	65	2 H72805	gpsi protein - Myc
38	43	8.5	74	2 H84256	hypothetical prote
39	43	8.5	76	1 IHER2	high potential iro
40	42.5	8.4	58	2 F84303	50S ribosomal prot
41	42.5	8.4	66	2 B83515	hypothetical prote
42	42.5	8.4	81	2 A41441	intestinal trefoil
43	42.5	8.4	82	2 S37118	class II histocomp
44	42	8.3	59	2 AF2245	hypothetical prote
45	42	8.3	62	2 A12767	hypothetical prote

ALIGNMENTS

RESULT 1

NTSR5L
C:Species: Leirus quinquestriatus (Egyptian scorpion)
C;Date: 30-Apr-1979 #sequence_revision 13-Jul-1981 #text_change 15-Oct-1996
C;Accession: A01741
R;Kopeyan, C.; Martinez, G.; Rochat, H.
FEBS Lett. 89, 54-58, 1978
A;Title: Amino acid sequence of neurotoxin V from the scorpion Leirus quinquestriatus
A;Reference number: A01741; MUID:78191225; PMID:658402
A;Note: L. q. quinquestriatus
A;Accession: A01741
A;Molecule type: Protein
A;Residues: 1-64 <KOP>
C;Superfamily: scorpion neurotoxin
C;Keywords: blocked carboxyl end; neurotoxin; venom
F;12-63,16-36,22-46,26-48/Disulfide bonds: #status Predicted
F;64/Modified site: blocked carboxyl end (Asn) (probably amidated) #status experiment

Query Match 10.7%; Score 54.5; DB 1; Length 64;
Best Local Similarity 34.0%; Pred. NO. 49;
Matches 16; Conservative 5; Mismatches 15; Indels 11; Gaps 3;

QY 35 RHSTFIPTGN-----KWGLEKNYCR--NPDGDLNGPWCTMNP 72
DB 10 KNTFFCGRNAYCNDECKKKGGSCGQWASPYG--NACWYKLPDR 54

RESULT 2

E82078
conserved hypothetical protein VC2429 [imported] - Vibrio cholerae (strain N16961 se
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: E82078
R;Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Seller:
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82078
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-65 <HEI>
A;Cross-references: GB:AE004312; GB:AE003852; NID:g9656995; PIDN:AAF95572.1; GSPDB:GI
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2429
A;Map position: 1

Query Match 10.7%; Score 54.5; DB 2; Length 65;
Best Local Similarity 31.7%; Pred. NO. 50;

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: G82600

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-68 <STM>

A;Cross-references: GB:AE004025; GB:AE003849; NID:G9107217; PIDN:AAF84897.1; GSPDB:GN

A;Experimental source: strain 9a5C

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.A.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to Genbank, June 2000

A;Authors: Ferreira, V.C.A.; Ferrio, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.H.; Miracca, E.C.; Miyaki, C. F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, R.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa M.; Teuhako da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L. A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF2098

Query Match 9.7%; Score 49.5; DB 2; Length 68;

Best Local Similarity 31.6%; Pred. No. 1.8e+02;

Matches 12; Conservative 3; Mismatches 6; Indels 17; Gaps 2;

QY 65 WCTMMPR-----KLFY-----CDIPLCAS 85

|||||

Db 5 WCYOGSTRAISELOELALRLLOAYAFMSFMSALPLCAS 42

|||||

RESULT 6

NTSREB

neurotoxin XI - scorpion (Buthus occitanus)

C;Species: Buthus occitanus tunetanus

C;Date: 17-May-1995 #sequence_revision 17-May-1985 #text_change 23-Aug-1996

C;Accession: A01746

R;Martin, M.F.; Rochat, H.

Toxicon 22, 279-291, 1984

A;Title: Purification of thirteen toxins active on mice from the venom of the North A

A;Reference number: A94316; MUID:84224814; PMID:6729843

A;Accession: A01746

A;Molecule type: protein

A;Residues: 1-65 <MAR>

C;Superfamily: scorpion neurotoxin

C;Keywords: neurotoxin; venom

F;12-63,16-36,22-46,26-48/Disulfide bonds: #status predicted

Query Match 9.5%; Score 48.5; DB 1; Length 65;

Best Local Similarity 28.6%; Pred. No. 2.2e+02;

Matches 12; Conservative 5; Mismatches 18; Indels 7; Gaps 1;

QY 35 RHSTFIPGNKWAGL-----EKYNPDGDPGNGPCYTM 69

|||||

Db 10 RNCYFGCTNAYCNBECYKLGESGYCQWVGRGNACWYKL 51

|||||

RESULT 7

PQ0852

glycoprotein H - human herpesvirus 6 (fragment)

C;Species: human herpesvirus 6

C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999

C;Accession: PQ0852; PQ0851; PQ0853

R;Gompels, J.A.; Carrigan, D.R.; Carss, A.L.; Arno, J.

J. Gen. Virol. 74, 613-622, 1993

A;Title: Two groups of human herpesvirus 6 identified by sequence analyses of laborat

A;Reference number: JQ2382; MUID:93224882; PMID:8385692

A;Accession: PQ0852

A;Molecule type: DNA

A;Residues: 1-65 <GOM>

A;Experimental source: strain C4

```

A:Accession: PQ0851
A:Molecule type: DNA
A:Residues: 1-65 <GOM1>
A:Experimental source: isolate C3
A:Accession: PQ0854
A:Molecule type: DNA
A:Residues: 1-65 <GOM2>
A:Experimental source: isolate C9
A:Accession: PQ0853
A:Molecule type: DNA
A:Residues: 1-65 <GOM2>
A:Experimental source: strain C6
C:Superfamily: herpesvirus glycoprotein H
C:Keywords: glycoprotein; membrane protein

Query Match          9.5%; Score 43.5; DB 2; Length 65;
Best Local Similarity 20.0%; Pred. NO. 2.2e+02;
Matches 14; Conservative 9; Mismatches 18; Indels 29; Gaps 4;

QY 23 TPCQEWAAQEPHRSHTFIPGKNWAGLEKNYCRNPDGDNPGWC-----YTMNPK 73
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 12 TPCYSW-----RPWTISNESCKN--GNSNPVIRPGITFNFTKNDTR 54

QY 74 LFDYCDIPLC 83
    : : : | |
Db 55 IV---QVQPK 61

RESULT 8
T35510
hypothetical protein SC6G10.01c - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A:Reference number: T21581
A:Accession: T35510
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-40 <SEE>
A:Cross-references: EMBL:AI049497; PIDN:CAB39855.1; GSPDB:GN000070; SCOEDB:SC6G10.01c
C:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC6G10.01c

Query Match          9.4%; Score 48; DB 2; Length 40;
Best Local Similarity 47.8%; Pred. NO. 1.5e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 4; Gaps 1;

QY 21 TCTPCQEWAAQEPHRSHTFIPGT 43
    : ||| : | | || | | | |
Db 7 SGTFRRAANPPPHR-----PCT 25

RESULT 9
S21332
hypothetical protein 2 - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 30-Jun-2001
R:Paulus, F.; Canaday, J.; Otten, L.
submitted to the EMBL Data Library, October 1990
A:Description: Limited host range Ti plasmids; recent origin from wide host range Ti pla
A:Reference number: S21331
A:Accession: S21332
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-59 <PAU>
A:Cross-references: EMBL:X55075; NID:g39138; PIDN:CAA38901.1; PID:g39140

Query Match          9.4%; Score 47.5; DB 2; Length 59;
Best Local Similarity 31.1%; Pred. NO. 2.6e+02;

```

```

Matches 14; Conservative 4; Mismatches 16; Indels 11; Gaps 2;

QY 22 GTPCQEWAAQEPHRSHTFIPGKNWAGLEKNYCRNPDGDNPGW 65
    | : | | : | | | | | | | | | | |
Db 16 GSPAQEPSSRRHS-----AGLLYRCRPRSSSPIHAAW 50

RESULT 10
D81186
hypothetical protein NME0542 [imported] - Neisseria meningitidis (strain MC58 serogr
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: D81186
R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Glin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:2015755; PMID:10710307
A:Accession: D81186
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <TET>
A:Cross-references: GB:AE002410; GB:AE002098; NID:g7225766; PIDN:AAF40971.1; PID:g72;
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NME0542

Query Match          9.4%; Score 47.5; DB 2; Length 67;
Best Local Similarity 29.0%; Pred. NO. 2.9e+02;
Matches 9; Conservative 5; Mismatches 14; Indels 3; Gaps 1;

QY 41 PGTNKNWAGLEKNY---CRNPDGDNPGWCYT 68
    || : | | : : | | : | | | |
Db 23 PGNQWNGHPRVFLPLCEGSGSVACPYCGT 53

RESULT 11
T6EPF5J
mambin - eastern Jameson's mamba
N:Alternate names: dendroaspin; short toxin S5C1
C:Species: Dendroaspis jamesoni kalmosae (eastern Jameson's mamba)
C:Date: 19-Feb-1984 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: A42561; S50910; A01683
R:McDowell, R.S.; Dennis, M.S.; Louie, A.; Shuster, M.; Mulkerin, M.G.; Lazarus, R.;
Biochemistry 31, 4766-4772, 1992
A:Title: Mambin, a potent glycoprotein IIB-IIIa antagonist and platelet aggregation :
A:Reference number: A42561; MUID:92273586; PMID:1591238
A:Accession: A42561
A:Molecule type: protein
A:Residues: 1-59 <WCD>
A:Cross-references: PIDN:AAB22225.1; PID:g249694
A:Experimental source: venom
R:Jaseja, M.; Lu, X.; Williams, J.A.; Sutcliffe, M.J.; Kakkar, V.V.; Parslow, R.A.; I
Eur. J. Biochem. 226, 861-868, 1994
A:Title: (1)H-NMR assignments and secondary structure of dendroaspin, an RGD-contain
A:Reference number: S50910; MUID:95112851; PMID:7813476
A:Accession: S50910
A:Molecule type: protein
A:Residues: 1-59 <JAS>
R:Joubert, F.J.; Taljaard, N.
Biochim. Biophys. Acta 579, 228-233, 1979
A:Title: Some properties and the complete primary structures of two reduced and S-ca
A:Reference number: A90630; MUID:79232598; PMID:465532
A:Accession: A01683
A:Molecule type: protein
A:Residues: 1-15, 'CT', 18-27, 'RNI', 28-32, 34-59 <JOU>
C:Comment: the cell attachment motif may serve in inhibiting platelet aggregation.
C:Superfamily: snake toxin
C:Keywords: toxin; venom
F:43-45/Region: cell attachment (R-G-D) motif

```


Result No.	Query	Score	Match	Length	DB	ID	Description
1	54.5	10.7	64	1	SCX5_LEIUOU	P01481 leivirus qui	
2	54.5	10.7	65	1	Y029_VIBCH	Q9xpel vibrio chol	
3	54.5	10.7	82	1	XC06_CONTE	Q9xzi1 conus texti	
4	52.5	10.3	64	1	SCX5_ADNA	P01482 androctonus	
5	49	9.6	67	1	GBGD_HUMAN	Q9p2w3 homo sapien	
6	49	9.6	67	1	GBGD_MOUSE	Q9jmf3 mus musculus	
7	48.5	9.5	64	1	YF29_VIBEA	Q87lt2 vibrio para	
8	48.5	9.5	65	1	SCXB_BTUTC	P01486 butus occi	
9	47	9.3	59	1	MAMB_DBNJA	P28375 dendroaspis	
10	47	9.3	65	1	SCBK_MESMA	P82815 mesobuthus	
11	46.5	9.2	85	1	SC15_MESMA	Q9ngx8 mesobuthus	
12	46.5	9.2	85	1	SCAL_MESMA	Q9ygz2 mesobuthus	
13	46	9.1	41	1	S482_HYSGI	P56854 hysteroocrat	
14	46	9.1	61	1	TX51_DBNJA	P01413 dendroaspis	
15	46	9.1	64	1	YG19_VIBVU	Q8ac30 vibrio vuln	
16	44	8.7	64	1	SCX3_LEIUOU	P01487 leivirus qui	
17	44	8.7	65	1	SIX1_BTUTC	P55902 butus occi	
18	44	8.7	74	1	TXC1_CUFGA	P81694 cupitennius	
19	44	8.7	85	1	SCXA_LEIQH	P17728 leivirus qui	
20	43.5	8.6	70	1	CSPH_SALTY	Q33793 salmonella	
21	43	8.5	65	1	VG51_BPMD2	O64241 mycobacteri	
22	43	8.5	68	1	NLT2_PRUAR	P82353 prunus arme	
23	43	8.5	76	1	HP12_ECTHA	P04169 ectothiorho	
24	42.5	8.4	58	1	RL37_HALNI	Q9hps3 halobacteri	
25	42.5	8.4	76	1	XC01_CONTE	Q9xzk8 conus texti	
26	42	8.3	64	1	IBB1_COILA	P07679 coix lachry	
27	42	8.3	76	1	SPV1_CHICK	Q9ptl1 gallus gall	
28	42	8.3	84	1	SCXB_MESMA	Q9n682 mesobuthus	
29	42	8.3	84	1	Y463_CAUCR	Q9axa9 caulobacter	
30	42	8.3	85	1	Y4X6_BACUS	Q04811 bacillus su	
31	41.5	8.2	49	1	R332_LISMO	Q9df31 listeria mo	
32	41.5	8.2	84	1	TFF1_HUMAN	P04155 homo sapien	
33	41	8.1	64	1	SCX4_MESMA	P58328 mesobuthus	

ID	YO29_VIBCH	STANDARD;	PRT;	63 AA.
20	correct.			

029_VIBCH	STANDARD.	PPT.	65 AA
0 VO20 VIBCH			

Y029_V1B0
TD V020

Y029_V1B0
TD V020

```

DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical zinc-binding UPF0243 protein VC2429.
GN VC2429.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=656;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20405833; PubMed=10953301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Hart D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -!- COPACTOR: Binds 1 zinc ion (By similarity).
CC -!- SIMILARITY: Belongs to the UPF0243 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF146359; AAD31919.1; -.
DR InterPro: IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
KW Toxin; Neurotoxin; Presynaptic neurotoxin; Ionic channel inhibitor;
KW Calcium channel inhibitor; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 50 OMEGA-TYPE CONOTOXIN TX06.
FT PEPTIDE 51 82 BY SIMILARITY.
FT DISULFID 53 71 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 70 81 BY SIMILARITY.
SQ SEQUENCE 82 AA; 9305 MW; 1739D2531B865860 CRC64;

Query Match 10.7%; Score 54.5; DB 1; Length 82;
Best Local Similarity 30.8%; Pred. No. 19;
Matches 12; Conservative 4; Mismatches 6; Indels 17; Gaps 1;

QY 47 AGLEKNCYCRNPDGNGWCTMPKLFYCDIPLCAS 85
Db | ||||| : ||| : |
46 AKLEKNCY-----QEKWDYCPVPFLGS 67

RESULT 4
SCX5_ANDMA
ID SCX5_ANDMA STANDARD; PRT; 64 AA.
AC P01482;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurotoxin V (AaH V) (AaH5).
OS Androctonus mauretanicus mauretanicus (Scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butoidae; Butidae; Androctonus.
OX NCBI_TaxID=6860;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=85193276; PubMed=3992595;
RA Rosso J.P., Kochat H.;
RT "Characterization of ten proteins from the venom of the Moroccan
RT scorpion Androctonus mauretanicus mauretanicus, six of which are
RT toxic to the mouse."
EL Toxicol 23:113-125(1985).
CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC of the activated channels, thereby blocking neuronal transmission.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.

```


ID	Y29_VTBPA	STANDARD;	PRT;	64 AA.
AC	Q87LF2;			
DC	15-SEP-2003 (Rel. 42, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
GN	Hypothetical zinc-binding UPF0243 protein VP2529.			
DE	VP2529.			
OS	Vibrio parahaemolyticus.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;			
OC	Vibrionaceae; Vibrrio.			
NCBI_TaxID=670;				
OR				
OR	[1]			
RP	SEQUENCE FROM N.A.			
RP	STRAN-RIMD 2210633 / Serotype O3:K6;			
RR	MEDLINE=2508454; PubMed=12620739;			
RA	Iijima K., Ohshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,			
RA	Makino Y., Nishima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,			
RA	Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;			
RT	"Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism			
RT	distinct from that of V. cholerae.";			
RT	Lancet 361:743-749(2003).			
RL	-I- COFACTOR: Binds 1 zinc ion (By similarity).			
CC	-I- SIMILARITY: Belongs to the UPF0243 family.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
CC	EMBL; AF005081; BAC60792.1; -			
DR	HMAP; MF_00649; -; 1.			
KK	Hypothetical protein; Zinc; Metal-binding; Complete proteome.			
FT	METAL 9			
FT	METAL 12 12			
FT	METAL 28 28			
FT	METAL 32 32			
FT	SEQUENCE 64 AA; 7187 MW; AD7939573E9FFEF9 CRG64;			
SQ				
Query Match 9.5%; Score 48.5; DB 1; Length 64;				
Best Local Similarity 29.0%; Pred. No. 69;				
Matches 18; Conservative 2; Mismatches 25; Indels 17; Gaps 3				
OY	16 KATTVGTGTC---QEWAAQEPHRSFTFGTK-----WAGLEKNYCRNPDGSDING 63			
DB	3 KITIVQCPQGQDVGEWGEQSPH-----PFCSKKKMTIDFGWADENAIAGAPDNMSD 57			
OY	64 PW 65			
DB	58 GW 59			
RESULT 8				
ID	SCXB_BUTOC	STANDARD;	PRT;	65 AA.
AC	P01486;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Neurotoxin XI.			
OS	Buthus occitanus tunetanus (Common European scorpion).			
OC	Enkaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;			
OC	Buthidae; Buthidae; Buthus.			
NCBI_TaxID=6871;				
OR				
OR	[1]			
RP	SEQUENCE.			
RP	TISSUE=Venom;			
RC	MEDLINE=84224814; PubMed=6729843;			
RA	Martin M.-F., Rochat H.;			
RT	"Purification of thirteen toxins active on mice from the venom of the			
RT	North African scorpion Buthus occitanus tunetanus.";			
RT				

```

CC THE BINDING OF PURIFIED PLATELET FIBRINOGEN RECEPTOR GPIIb-IIIa
CC TO IMMOBILIZED FIBRINOGEN.
CC
CC -!- SUBCELLULAR LOCATION: Secreted.
CC
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC
CC -!- SIMILARITY: belongs to the snake toxin family.
CC
DR PIR: A42561, P6EP5J.
DR
DR PDB: 1DRS; 20-DEC-94.
DR
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; toxin; 1.
DR
DR ProDom: PD00206; Snake_toxin; 1.
DR
DR PROSITE: PS00272; SNAKE_TOXIN; FALSE_NEG.
KW Blood coagulation; Cell adhesion; Toxin; 3D-structure.
FT DISULFID 3 22
FT DISULFID 17 37
FT DISULFID 39 51
FT DISULFID 52 57
FT SITE 43 45
FT STRAND 2 4
FT TURN 8 9
FT STRAND 14 16
FT STRAND 22 24
FT TURN 28 31
FT STRAND 36 37
FT STRAND 50 52
FT SEQUENCE 59 AA; 6754 MW; 20DDC6A5D9DFE41 CRC64;
CC
Query Match 9.3%; Score 47; DB 1; Length 59;
Best Local Similarity 30.5%; Pred.No.93;
Matches 16; Conservative 3; Mismatches 14; Indels 24; Gaps 4;
CC
QY 15 KKATVTGTPCQE-----WAAQEPHRHSTIPGNTKWAGLEKNCRNPDGSDINGPWC 66
Db 10 KPPTET--CQEDSCYKNIW-----TFDNIIRGGG----CFWPRGDMGPYC 51
CC
RESULT 10
ID SCBK_MESMA STANDARD; PRT; 65 AA.
AC P82815;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bukatoxin.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OX Buthidea; Buthidae; Mesobuthus.
OX NCBI_taxID=34649;
RN [1]
RP SEQUENCE.
RC TISSUE=venom;
RX MEDLINE=21211325; PubMed=11311230;
RA Srinivasan K.N., Nirthanan S., Sasaki T., Sato K., Cheng B.,
RA Gwee M.C.E., Kini R.M., Gopalakrishnakone P.;
RA "Functional site of bukatoxin, an alpha-type sodium channel
RA neurotoxin from the Chinese scorpion (Buthus martensii Karsch) venom:
RA probable role of the (52)PKVP(56) loop.";
RL FEBS Lett. 494:145-149(2001).
CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC of the activated channels, thereby blocking neuronal transmission.
CC
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC
CC -!- ALPHA-TOXIN SUBFAMILY.
CC
DR HSP: P17728; ILQ1.
DR
DR InterPro: IPR003614; Kntcl.
DR
DR InterPro: IPR001219; Neurotoxin.
DR
DR InterPro: IPR002061; scorpion_toxinL.
DR
DR Pfam: PF00537; toxin_3; 1.
DR
DR PRINTS: PR00284; TOXIN.
DR
DR ProDom: PD000908; Scorpion_toxinL; 1.
DR
DR SMART: SM00505; Kntcl; 1.
DR
DR Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor; Toxin.
KW DISULFID 12 63
KW BY SIMILARITY

```

```

FT DISULFID 16 36 BY SIMILARITY.
FT DISULFID 22 46 BY SIMILARITY.
FT DISULFID 26 48 BY SIMILARITY.
SQ SEQUENCE 65 AA; 7226 MW; F2AYD760E7B259DA CRC64;

Query Match 9.3%; Score 47; DB 1; Length 65;
Best Local Similarity 25.5%; Pred. No. 1e+02;
Matches 12; Conservative 7; Mismatches 20; Indels 8; Gaps 2;

QY 35 RHSTFTPGINKWA-----GLKNYCRNPDGDPGNGWPCYTMNPKL 74
DQ 10 KNCAYFCGRNAYCDECIINGAESGYCQAGYVGNACWCYKL-PDKV 55

RESULT 11
SC15_MESMA
ID SC15_MESMA STANDARD; PRT; 85 AA.
AC Q9GNG8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-neurotoxin TX15 precursor.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=venom gland;
RX MEDLINE=20317244; PubMed=10858508;
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Liu H., Jiang D.-H., Mao X.;
RT "Nine novel precursors of Buthus martensii scorpion alpha-toxin
homologues."
RL Toxinon 38:1653-1661(2000).
CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
of the activated channels, thereby blocking neuronal transmission.
CC This toxin is active against mammals (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
ALPHA-TOXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF163017; AAC39643.1; -.
DR EMBL; AF163016; AAC39642.1; -.
DR HSP; P17728; ILQI.
DR InterPro; IPR003614; Knot1.
DR InterPro; IPR001219; Neurotoxin.
DR InterPro; IPR002061; Scorpion_toxinL.
DR Pfam; PF00537; toxin_3; 1.
DR PRINTS; PR00284; TOXIN.
DR ProDom; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Knot1; 1.
DR Toin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW Signal.
FT SIGNAL. 1 19 POTENTIAL.
FT CHAIN 20 85 ALPHA-NEUROTOXIN TX15.
FT DISULFID 31 82 BY SIMILARITY.
FT DISULFID 35 55 BY SIMILARITY.
FT DISULFID 41 65 BY SIMILARITY.
FT DISULFID 45 67 BY SIMILARITY.
SQ SEQUENCE 85 AA; 9398 MW; 3312B422E1C57E1 CRC64;

Query Match 9.2%; Score 46.5; DB 1; Length 85;
Best Local Similarity 33.3%; Pred. No. 1.6e+02;
Matches 11; Conservative 4; Mismatches 7; Indels 11; Gaps 2;

QY 42 GTKNWAGLEKYNCRNPDGDPGNGWPCYTMNPKL 74
DQ 53 GYCQWAGVYGNAC-----WCYKL-PDKV 74

RESULT 13
S482_HYSGI
ID S482_HYSGI STANDARD; PRT; 41 AA.
AC P56854;

```

```

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Toxin SNX-482.
OS Hysterochrates gigas (African tarantula).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Mygalomorphae; Theraphosidae; Hysterochrates.
OX NCBI_TaxID=118972;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=99017889; PubMed=9799496;
RA Newcomb R., Szoke B., Palma A., Wang G., Chen X.H., Hopkins W.,
RA Cong R., Miller J., Urge L., Parczyk-Hornoch K., Loo J.A., Dooley D.J.,
RA Radasdi L., Tsien R.W., Lemos J., Miljanich G.;
RT "selective peptide antagonist of the class E calcium channel from the
RT venom of the tarantula Hysterochrates gigas.";
RL Biochemistry 37:15353-15362(1998).
CC -!- FUNCTION: POTENT AND SELECTIVE BLOCKER OF THE CLASS E CALCIUM
CC CHANNEL.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE SPIDER POTASSIUM CHANNEL INHIBITORY
CC TOXIN FAMILY.
DR HSP; P56852; 1D1H.
KW Toxin; Neurotoxin; Ionic channel inhibitor; Calcium channel inhibitor.
SQ SEQUENCE 41 AA; 4501 MW; C5D3CC96D3A0CEFF4 CRC64;

Query Match 9.1%; Score 46; DB 1; Length 41;
Best Local Similarity 33.3%; Pred. No. 83;
Matches 11; Conservative 4; Mismatches 16; Indels 2; Gaps 1;

QY 48 GLEKYNCRNPDG-DINGPCWYTMNPKLPDYC 78
|::| | | | | | | | | | | | | | | |
DB 1 GVDKAGCYMFGGSVNDCCPRLGCHSLFSYC 33

RESULT 14
TX51_DENJA
ID TX51_DENJA STANDARD; PRT; 61 AA.
AC P01413;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Toxin S5C1.
OS Dendroaspis jamesoni kaimosae (Eastern Jameson's mamba).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosaurs; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Elapinae; Dendroaspis.
OX NCBI_TaxID=8619;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=79232598; PubMed=465532;
RA Joubert F.O., Taljaard N.;
RT "Some properties and the complete primary structures of two reduced
RT and S-carboxymethylated polypeptides (S5C1 and S5C10) from
RT Dendroaspis jamesoni kaimosae (Jameson's mamba) venom.";
RL Biochim. Biophys. Acta 579:228-233(1979).
CC -!- FUNCTION: INHIBITS ADP-INDUCED PLATELET AGGREGATION AND INHIBITS
CC THE BINDING OF PURIFIED PLATELET FIBRINOGEN RECEPTOR GPIIb-IIIa
CC TO IMMOBILIZED FIBRINOGEN (BY SIMILARITY TO NAMBIN).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: Belongs to the snake toxin family.
DR HSP; P28375; 1DRS.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00067; toxin; 1.
DR PRODOM: PD000206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; FALSE_NEG.
KW Blood coagulation; Cell adhesion; Toxin.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 16 39 BY SIMILARITY.

```

```

FT DISULFID 41 53 BY SIMILARITY.
FT DISULFID 54 59 BY SIMILARITY.
FT SITE 45 47 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 61 AA; 7024 MW; 34B25A68D48F7EAA CRC64;

Query Match 9.1%; Score 46; DB 1; Length 61;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 54 CRNPDGDIINGPMC 66
| | | | | | | | | | | | | | | |
DB 41 CFTPRGDMPEPYC 53

RESULT 15
YG19_VIBVU
ID YG19_VIBVU STANDARD; PRT; 64 AA.
AC Q8DC30;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical zinc-binding UPF0243 protein VV11619.
GN VV11619.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: Binds 1 zinc ion (By similarity).
CC -!- SIMILARITY: Belongs to the UPF0243 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF016802; AAC10038.1; ALT_INIT.
DR HMAP; MF_00649; -, 1.
KW Hypothetical protein; Zinc; Metal-binding; Complete proteome.
FT METAL 9 9 ZINC (BY SIMILARITY).
FT METAL 12 12 ZINC (BY SIMILARITY).
FT METAL 28 28 ZINC (BY SIMILARITY).
FT METAL 32 32 ZINC (BY SIMILARITY).
SQ SEQUENCE 64 AA; 7240 MW; AB06E7809FF59FE CRC64;

Query Match 9.1%; Score 46; DB 1; Length 64;
Best Local Similarity 27.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 3; Mismatches 23; Indels 20; Gaps 3;

QY 18 TWTGTFQC-----EWAAQEPHRHSTFPGTNK-----WAGLEKYNCRNPDGIN 62
| | | | | | | | | | | | | | | | | | | | | |
DB 2 TKTIIVCPCQGTNWVGEQSPHR-----PFCSKKQCIDFGWEADENAIIPGAPMDS 56
| | | | | | | | | | | | | | | | | | | | | |
QY 63 GFW 65
|
DB 57 DGW 59

Search completed: September 29, 2003, 08:47:29
Job time : 24 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:44:01; Search time 94 seconds

(without alignments)
236.091 Million cell updates/sec

Title: US-10-088-548-8

Perfect score: 508

Sequence: 1 EDCMFGNGKGYRKKATV.....YTMNPKLFYCDIPLCASS 86

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 105943

Minimum DB seq length: 0

Maximum DB seq length: 86

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL_23.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.podent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriap.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130	25.6	75	6 Q9BGN9	Q9bgn9 bos taurus
2	127.5	25.1	60	4 Q9UKJ7	Q9ukj7 homo sapien
3	123	24.2	53	4 Q9UMI2	Q9umi2 homo sapien
4	56	11.0	67	2 Q9FOV6	Q9fov6 azoarcus sp
5	54.5	10.7	65	16 Q9KPE1	Q9kpe1 vibrio chol
6	54	10.6	25	4 Q9UD88	Q9ud88 homo sapien
7	53	10.4	84	5 Q9W0Y8	Q9w0y8 drosophila
8	51.5	10.1	71	5 Q8MQD3	Q8mqd3 caenorhabdi
9	51	10.0	65	12 Q55745	Q55745 chilo iride
10	50	9.8	48	3 Q8TGR4	Q8tgr4 saccharomyc
11	49.5	9.7	51	10 Q8GRV2	Q8grv2 oryza sativ
12	49.5	9.7	68	16 Q9PBP2	Q9pbp2 xyella fas
13	48.5	9.5	61	15 Q89143	Q89143 visna virus
14	48.5	9.5	85	5 Q9ELM4	Q9elm4 androctonus
15	48.5	9.5	86	15 Q8ADD0	Q8add0 human immun
16	48	9.4	38	5 Q8WP57	Q8wp57 lutzomyia l

17	48	9.4	38	5 Q6WRR7	Q6wrr7 lutzomyia u
18	48	9.4	38	5 Q6WRR5	Q6wrr5 lutzomyia e
19	48	9.4	38	5 Q6WRR9	Q6wrr9 lutzomyia i
20	48	9.4	38	5 Q6WRR2	Q6wrr2 lutzomyia m
21	48	9.4	38	5 Q6WRR4	Q6wrr4 lutzomyia l
22	48	9.4	38	5 Q6WRR3	Q6wrr3 lutzomyia d
23	48	9.4	38	5 Q6WRR8	Q6wrr8 lutzomyia w
24	48	9.4	38	16 Q8XW59	Q8xw59 ralstonia s
25	47.5	9.4	28	3 Q9URD3	Q9urd3 aspergillus
26	47.5	9.4	59	2 Q44448	Q44448 agrobacteri
27	47.5	9.4	67	16 Q9KOP3	Q9kop3 neisseria m
28	47.5	9.4	78	15 Q993N1	Q993n1 human immun
29	47	9.3	60	15 Q89146	Q89146 visna virus
30	47	9.3	64	15 Q89145	Q89145 visna virus
31	47	9.3	67	11 Q9ER13	Q9er13 rattus norv
32	47	9.3	77	16 Q8YK17	Q8yk17 anabaena sp
33	47	9.3	86	10 Q9LQ96	Q9lq96 oryza sativ
34	47	9.3	86	11 Q8CJA5	Q8cja5 mus musculu
35	46.5	9.2	51	16 Q8YHK9	Q8yhk9 brucella me
36	46.5	9.2	62	17 Q26228	Q26228 methanobact
37	46.5	9.2	65	15 Q89144	Q89144 visna virus
38	46.5	9.2	73	16 Q9KD21	Q9kd21 bacillus ha
39	46.5	9.2	85	5 Q9GUA7	Q9gua7 mesobuthus
40	46.5	9.2	85	5 Q9SP69	Q9sp69 mesobuthus
41	46.5	9.2	86	4 Q9UFL9	Q9ufl9 homo sapien
42	46.5	9.2	86	15 Q8AD74	Q8ad74 human immun
43	46	9.1	71	12 Q85298	Q85298 orf virus.
44	46	9.1	75	11 Q8R043	Q8r043 mus musculu
45	45.5	9.0	52	12 Q92320	Q92320 human respi

ALIGNMENTS

RESULT 1

ID	Q9BGN9	PRELIMINARY;	PRT;	75 AA.
AC	Q9BGN9;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Hepatocyte growth factor (Fragment).			
GN	HGF.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Uterus;			
RA	Murakami S., Fujiwara C., Miyamoto Y., Takeuchi S., Takahashi S.,			
RA	Okuda K.;			
RT	"Expression and action of hepatocyte growth factor in bovine			
RT	endometrial stromal and epithelial cells in vitro.";			
RL	Submitted (PEB-2001) to the EMBL/GenBank/DBJ databases.			
CC	!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.			
DR	EMBL; AB056447; BAB33031.1; -.			
DR	HSSP; P14210; LBHT.			
DR	InterPro; IPR000001; Kringle.			
DR	Pfam; PF00051; Kringle; 2.			
DR	PRINTS; PR00018; KRINGLE.			
DR	ProDom; PD000395; Kringle; 2.			
DR	SMART; SM00130; KR; 2.			
DR	PROSITE; PS00021; KRINGLE_1; 1.			
DR	PROSITE; PS50070; KRINGLE_2; 2.			
KW	Glycoprotein; Kringle.			
FT	NON_TER 1			
FT	NON_TER 75			
FT	NON_TER 75			
SQ	SEQUENCE 75 AA; 8831 MW; 829EEFFCC49701B1 CRC64;			
Query Match		25.6%;	Score 130;	DB 6;
Best Local Similarity		57.6%;	Pred. No. 1.3e-07;	Length 75;

```

Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 52 NYCNRPDGSDINGPWCYTMNPKLFYCDIPLCA 84
    ||||| 1: ||||| 1: : ||||| 1:
Db 1 NYCNRPDGEGGFWCTSNPEVRYCVDIPLQCS 33

RESULT 2
Q9UKJ7 PRELIMINARY; PRT; 60 AA.
AC Q9UKJ7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=21181705; PubMed=11285247;
RA Ogorekova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringle IV
RT types 6 to 10 domain affect lip(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158657; AAF03676.1; -.
DR HSSP; P00747; IKRN.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6799 MW; 5719AA36B3E0FFD CRC64;

Query Match 25.1%; Score 127.5; DB 4; Length 60;
Best Local Similarity 55.6%; Pred. No. 2e-07;
Matches 20; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 48 GLENYCRNPDGSDINGPWCYTMNPKLFYCDIPLC 83
    || ||||| 1: ||||| 1: : ||||| 1:
Db 1 GLTNYCRNPDGEIR-PWCYTMNPSVWEYCNLTQC 35

RESULT 3
Q9UMI2 PRELIMINARY; PRT; 53 AA.
AC Q9UMI2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE .PIG protein (fragment).
GN PIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; K02921; AAA60123.1; -.

```

```

DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
FT NON_TER 1
FT NON_TER 53
SQ SEQUENCE 53 AA; 5893 MW; 2EB36E715CD40A47 CRC64;

Query Match 24.2%; Score 123; DB 4; Length 53;
Best Local Similarity 54.3%; Pred. No. 5,9e-07;
Matches 19; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 QDCMFGNGKGYRGKATVTGTPCQENWAAQPHRH 36
    |||| 1: |||| 1: |||| 1: |||| 1:
Db 9 QDCYHGDQSYRGTSSTTTTGKKCQSWSSMTPHRH 43

RESULT 4
Q9F0V6 PRELIMINARY; PRT; 67 AA.
AC Q9F0V6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Electron carrier protein ferredoxin N.
GN FDXN.
OS Azoarcus sp. BH72.
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Azoarcus.
OX NCBI_TaxID=62928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BH72;
RX MEDLINE=21264388; PubMed=11371540;
RA Egner T., Martin D.E., Sarkar A., Reinhold-Hurek B.;
RT "Role of a Ferredoxin Gene Cotranscribed with the nifHDK Operon in
RT N(2) Fixation and Nitrogenase 'Switch-Off' of Azoarcus sp. Strain
RT BH72.";
RL J. Bacteriol. 183:3752-3760(2001).
DR EMBL; AF200742; AAG35589.1; -.
DR HSSP; P00193; IDUR.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR Pfam; PF00037; fer4; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW 4Fe-4S; Iron; Iron-sulfur.
SQ SEQUENCE 67 AA; 6997 MW; 8E5A0E4A6FPE157B CRC64;

Query Match 11.0%; Score 56; DB 2; Length 67;
Best Local Similarity 33.3%; Pred. No. 44;
Matches 12; Conservative 6; Mismatches 16; Indels 2; Gaps 1;

QY 49 LEKNYCRNPDGSDINGPWCYTMNPKLFYCDIPLCA 84
    : 1: |||| 1: |||| 1: |||| 1:
Db 34 INKDICTCEGDFDEPQCLAVCPAG--DACIVPLAA 67

RESULT 5
Q9KPE1 PRELIMINARY; PRT; 65 AA.
AC Q9KPE1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein VC2429.
GN VC2429.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;

```


RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
 clone:PO438G07."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
 clone:B1126F07."
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP004398; BAC22372.1; -
 DR EMBL; AP005674; BAC22483.1; -
 SQ SEQUENCE 51 AA; 5702 MW; B4EC539E549BFF6C CRC64;

 Query Match 9.7%; Score 49.5; DB 10; Length 51;
 Best Local Similarity 29.4%; Pred. No. 1.8e+02;
 Matches 15; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

 QY 5 MFGNGKGVGKATVGTGTCQWAAQEPHRSSTFPGTNKAGLEKNVCR 55
 DB 1 MGEGRRRRRGEGEEIGGE-EEERGSEGRGDEIDPMLSKYPGEESCR 50

 RESULT 12
 Q9PBP2
 ID Q9PBP2 PRELIMINARY; PRT; 68 AA.
 AC Q9PBP2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein Xf2098.
 GN Xf2098.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Canargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva P.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa."
 RL Nature 406:151-159(2000).
 DR EMBL; AE004025; AAF84897.1; -

KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 68 AA; 7349 MW; 87902D67E5522F41 CRC64;

 Query Match 9.7%; Score 49.5; DB 16; Length 68;
 Best Local Similarity 31.6%; Pred. No. 2.5e+02;
 Matches 12; Conservative 3; Mismatches 6; Indels 17; Gaps 2;

 QY 65 WCYTNMPR-----KLFDY-----CDIPLCAS 85
 DB 5 WCIQSTRAAISLQELALRLLOAAAFMSALPLCAS 42

 RESULT 13
 Q89143
 ID Q89143 PRELIMINARY; PRT; 61 AA.
 AC Q89143;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Envelope polyprotein (Fragment).
 GN ENV.
 OS Yisna virus.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=676;
 RX MEDLINE=95302052; PubMed=7782784;
 RA Leroux C., Vuillemoz S., Mornex J.F., Greenland T.;
 RT "Genomic heterogeneity in the pol region of ovine lentiviruses
 obtained from bronchoalveolar cells of infected sheep from France."
 RL J. Gen. Virol. 76:1533-1537(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=676;
 RX MEDLINE=97372948; PubMed=9229003;
 RA Leroux C., Chastang J., Greenland T., Mornex J.F.;
 RT "Genomic heterogeneity of small ruminant lentiviruses: existence of
 heterogeneous populations in sheep and of the same lentiviral
 genotypes in sheep and goats."
 RL Arch. Virol. 142:1125-1137(1997).
 DR EMBL; U35858; AAA78287.1; -
 KW Polyprotein.
 FT NON_TER 1
 FT NON_TER 61
 SQ SEQUENCE 61 AA; 7240 MW; C88726FF6FB078B0 CRC64;

 Query Match 9.5%; Score 48.5; DB 15; Length 61;
 Best Local Similarity 26.5%; Pred. No. 2.9e+02;
 Matches 13; Conservative 7; Mismatches 18; Indels 11; Gaps 2;

 QY 26 QEWA-----QEPHRSSTFPGTNKAGLEKNY-CRNPDDG 63
 DB 13 QQWTCARVKKWKIKQELKRDLSLYIAGDFWGVKAKYSCRNIGLDG 61

 RESULT 14
 Q9BLM4
 ID Q9BLM4 PRELIMINARY; PRT; 85 AA.
 AC Q9BLM4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative toxin precursor.
 OS Androctonus australis (Sahara scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthoidea; Buthidae; Androctonus.
 OX NCBI_TaxID=6858;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21211344; PubMed=11311249;
 RA Ceard B., Martin-Eauclaire M.F., Bougis P.E.;
 RT "Evidence for a position-specific deletion as an evolutionary link

Search completed: September 29, 2003, 08:49:13
Job time : 98 secs